

UNIVERSITATEA NAȚIONALĂ DE ȘTIINȚĂ ȘI TEHNOLOGIE
POLITEHNICA DIN BUCUREȘTI
FACULTATEA DE ELECTRONICĂ, TELECOMUNICAȚII
ȘI TEHNOLOGIA INFORMAȚIEI



LUCRARE DE DIPLOMĂ

Identificarea stadiilor de somn la copii și adolescenți
prin metode automate de analiză a semnalelor EEG

Coordonatori științifici:

Conf. dr. ing. Horia CUCU

S.L. Dr. Ana-Antonia NICOLAE

Absolvent:

Erol-Andrei TOMA

BUCUREȘTI, 2026

**NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY
POLITEHNICA BUCHAREST
FACULTY OF ELECTRONICS, TELECOMMUNICATIONS
AND INFORMATION TECHNOLOGY**



BACHELOR THESIS

**Automatic Sleep Stage Classification in Children
and Adolescents Using EEG Signal Analysis**

Thesis advisors:

Conf. dr. ing. Horia CUCU

S.L. Dr. Ana-Antonia NICOLAE

Graduate:

Erol-Andrei TOMA

BUCHAREST, 2026

Declarație de onestitate academică

Prin prezent declar că lucrarea cu titlul *Identificarea stadiilor de somn la copii și adolescenți prin metode automate de analiză a semnalelor EEG*, prezentată în cadrul Facultății de Electronică, Telecomunicații și Tehnologia Informației a Universității “Politehnica” din București ca cerință parțială pentru obținerea titlului de Inginer în domeniul Inginerie electronică, telecomunicații și tehnologii informaționale, programul de studii Microelectronică, optoelectronică și nanotehnologii, este scrisă de mine și nu a mai fost prezentată niciodată la o facultate sau instituție de învățământ superior din țară sau străinătate. Declar că toate sursele utilizate, inclusiv cele de pe Internet, sunt indicate în lucrare, ca referințe bibliografice. Fragmentele de text din alte surse, reproduse exact, chiar și în traducere proprie din altă limbă, sunt scrise între ghilimele și fac referință la sursă. Reformularea în cuvinte proprii a textelor scrise de către alți autori face referință la sursă. Înțeleg că plagiatul constituie infracțiune și se sancționează conform legilor în vigoare. Declar că toate rezultatele simulărilor, experimentelor și măsurărilor pe care le prezint ca fiind făcute de mine, precum și metodele prin care au fost obținute, sunt reale și provin din respectivele simulări, experimente și măsurători. Înțeleg că falsificarea datelor și rezultatelor constituie fraudă și se sancționează conform regulamentelor în vigoare.

București, Iunie 2026.

Absolvent: Erol-Andrei TOMA



Cuprins

Sinopsis	5
Abstract	6
1 Introducere	7
1.1 Context	7
1.2 Problema	7
1.3 Obiective	8
2 Baze teoretice	10
2.1 Biologia somnului	10
2.1.1 Structura temporală a somnului normal	10
2.2 Stadiile de somn și EEG	11
2.2.1 Principiile înregistrării EEG	11
2.2.2 Benzile de frecvență EEG relevante pentru somn	11
2.2.3 Stadiile de somn conform manualului AASM	12
2.2.4 Particularitățile EEG pediatric	13
3 State of the Art	14
3.1 DeepSleepNet	14
3.1.1 Arhitectura modelului	14
3.1.2 Limitări	16
3.2 AttnSleep	16
3.2.1 Arhitectura modelului	16
3.2.2 Limitări	18
3.3 PediaSleepNet (modele proprii)	18
3.3.1 PediaSleepNet Base	19
3.3.2 PediaSleepNet V2	21
3.3.3 Limitări	22
3.4 Sintează comparativă	23
4 Seturi de date	24
4.1 Sleep-EDF	24

4.2	HMC Sleep Staging	24
4.3	Setul de date pediatric	25
5	Metodologie	28
5.1	Preprocesarea semnalelor	28
5.1.1	Preprocesarea fiecărui dataset	28
5.2	Antrenarea modelelor (pre-antrenare pe date adulte)	29
5.2.1	DeepSleepNet	29
5.2.2	AttnSleep	30
5.2.3	PediaSleepNet	30
5.3	<i>Adaptarea</i> pe date pediatrice	31
5.3.1	Strategia de <i>adaptare</i>	32
5.4	Mediul software și hardware	33
5.5	Metrici de evaluare	34
5.5.1	Acuratețe globală (<i>Accuracy</i>)	34
5.5.2	<i>F1-score</i> per clasă	35
5.5.3	<i>Macro F1-score</i>	35
5.5.4	Cohen's κ	35
5.5.5	Matricea de confuzie	36
6	Rezultate și evaluare	37
6.1	Rezultate la pre-antrenare	37
6.1.1	DeepSleepNet (Sleep-EDF)	37
6.1.2	DeepSleepNet (HMC)	39
6.1.3	AttnSleep (Sleep-EDF)	40
6.1.4	AttnSleep (HMC)	41
6.1.5	PediaSleepNet Base (Sleep-EDF)	42
6.1.6	PediaSleepNet Base (HMC)	43
6.1.7	PediaSleepNet V2 (Sleep-EDF)	44
6.1.8	PediaSleepNet V2 (HMC)	45
6.2	Rezultate la <i>adaptarea</i> pe date pediatrice	45
6.2.1	DeepSleepNet (Sleep-EDF - pediatric)	46
6.2.2	DeepSleepNet (HMC - pediatric)	47
6.2.3	AttnSleep (Sleep-EDF - pediatric)	48
6.2.4	AttnSleep (HMC - pediatric)	49
6.2.5	PediaSleepNet Base (Sleep-EDF - pediatric)	50
6.2.6	PediaSleepNet Base (HMC - pediatric)	51
6.2.7	PediaSleepNet V2 (Sleep-EDF - pediatric)	52
6.2.8	PediaSleepNet V2 (HMC - pediatric)	53

7	Concluzii	54
7.1	Concluzii privind pre-antrenarea pe date adulte	54
7.2	Concluzii privind transferul pe date pediatrice	56
7.3	Limitele cercetării	56
7.4	Contribuții proprii	57
7.5	direcții viitoare de cercetare	57
	Bibliografie	59
A	Anexe	
A.1	Configurația și hiperparametrii modelelor	62
A.2	Setul de date pediatric	62

Sinopsis

Monitorizarea clinică a somnului se realizează prin polisomnografie (PSG) combinată cu analiza semnalului electroencefalografic (EEG), acestea constituind instrumentele de referință în diagnosticul tulburărilor de somn. Clasificarea manuală a înregistrărilor PSG rămâne un proces consumator de timp și afectat de diferențe între evaluatori de până la 20% în cazul stadiilor ambigue [12], fapt care a împins cercetarea înspre a căuta metode automate de clasificare.

Prezenta lucrare abordează problema identificării automate a stadiilor de somn la o populație care este mai puțin studiată, copii și adolescenți, pornind de la semnale EEG monocanal și arhitecturi de tip deep learning. Pacienții din această grupă de vârstă prezintă caracteristici EEG bine documentate în neurofiziologie [3]: amplitudini semnificativ mai mari față de adulți, un spectru dominant deplasat spre frecvențe mai joase și o arhitectură a somnului care se modifică în paralel cu maturizarea cerebrală. Drept consecință, modelele calibrate pe date de adulți produc erori atunci când sunt aplicate la copii, iar literatura de specialitate dedicată acestei populații este net mai restrânsă față de cea pentru adulți.

Lucrarea evaluează comparativ patru arhitecturi: **DeepSleepNet** [25], **AttnSleep** [6], **PediaSleepNet Base** și **PediaSleepNet V2**. Ultimele două fiind arhitecturi originale propuse în această lucrare, cu filtrare adaptivă pe vârstă și condiționare prin mecanismul FiLM [21], iar V2 cu un cap auxiliar dedicat detecției stadiului N1. Fiecare model este mai întâi antrenat pe seturi de date publice de adulți Sleep-EDF [7] și HMC [1] după care este adaptat prin *fine-tuning* pe un set de date pediatric. Aportul central al lucrării constă într-o analiză cantitativă sistematică a transferului între domenii adult->pediatric și în propunerea PediaSleepNet ca răspuns arhitectural explicit la semnalul specific EEG la vârste de dezvoltare.

Abstract

Clinical evaluation of sleep relies on polysomnography (PSG) coupled with electroencephalographic (EEG) signal analysis, which together form the diagnostic backbone for a wide spectrum of sleep disorders. Scoring PSG recordings by hand is both time-intensive and subject to inter-rater disagreement reaching 20% on ambiguous stages [12], which has driven considerable research interest toward automated staging approaches.

The work presented here focuses on deep learning-based automatic sleep stage identification, in children and adolescents, a cohort that has received far less attention than the adult population in the automated sleep staging literature. Children’s EEG signals differ from those of adults in well-established ways: higher amplitudes, lower dominant frequencies, and a sleep architecture that shifts substantially as the brain matures [3]. These characteristics mean that models trained on adult recordings tend to perform poorly when applied to younger subjects, yet systematic comparisons of architectures specifically for this population remain scarce.

Three architectures are compared head-to-head: **DeepSleepNet** [25], **AttnSleep** [6], and **PediaSleepNet**. The last two are novel architectures developed in this work, integrating age conditioning through Feature-wise Linear Modulation, FiLM [21]. All four models are first trained on publicly available adult datasets Sleep-EDF [7] and HMC [1] and then fine-tuned on a pediatric dataset. The central contribution of the thesis is a quantitative, systematic examination of the limitations inherent to adult->pediatric cross-domain transfer, alongside the design and evaluation of PediaSleepNet as an architecture built around the developmental characteristics of pediatric EEG signals.

Capitolul 1

Introducere

1.1 Context

Aproximativ o treime din viața fiecărei persoane este petrecută dormind, ceea ce arată importanța somnului în biologia umană. Somnul nu este o stare pasivă: el coordonează procese esențiale precum consolidarea memoriei, secreția hormonului de creștere, regenerarea celulară și maturizarea progresivă a sistemului nervos central, în special în copilărie și adolescență [3]. Perturbarea sau fragmentarea somnului rezultă în deficiente cognitive, dereglare metabolică și risc cardiovascular crescut pe termen lung.

Metoda de referință pentru evaluarea obiectivă a somnului este prin prisma **poli-somnografiei** (PSG): înregistrarea simultană, pe durata unei nopți, a mai multor semnale fiziologice precum electroencefalograma (EEG), electrooculograma (EOG), electromiograma mentonului (EMG), electrocardiograma (ECG) și, când este necesar, parametri respiratori. Clinicienii utilizează aceste semnale pentru a clasifica somnul în stadii distincte, conform manualului publicat de American Academy of Sleep Medicine (AASM) [12]. O singură înregistrare nocturnă produce câteva sute de epoci (epochs) de câte 30 de secunde, fiecare dintre ele trebuind clasificate individual.

Datorită creșterii volumului de date PSG digitizate și al avansurilor metodelor de *deep learning*, clasificarea automată a stadiilor de somn a evoluat rapid într-un domeniu de cercetare activ. Primele arhitecturi end-to-end au demonstrat că rețelele convoluționale [17] și recurente pot extrage direct din semnalul EEG brut reprezentări suficient de discriminative pentru a egala sau aproxima performanța clasificării umane pe date de adulți [25, 6]. Cu toate acestea, aproape întreg acest mod de lucru a folosit exclusiv populații adulte, ignorând o problemă clinică distinctă: specificitatea semnalului EEG la vârste de dezvoltare.

1.2 Problema

Aplicarea clasificării automate a stadiilor de somn la copii și adolescenți nu este o simplă extindere a cazului adult.

Profilul EEG pediatric diferă structural de cel adult. Maturizarea cortexului cerebral, un proces care se desfășoară pe tot parcursul copilăriei și se termină abia în adolescența târzie, se reflectă direct în semnalul EEG înregistrat [3]. Undele delta din stadiul N3 sunt intrinsec de amplitudine mai mare la copii, activitatea theta din N2 este mai pronunțată, iar frecvența de fond este deplasată spre valori mai mici față de cea observată la adulți. Frecvența fusurilor de somn (*sleep spindles*), care marchează clasic stadiul N2, este ușor mai ridicată la copii (aproximativ 13–15 Hz) față de intervalul de 12–14 Hz raportat la adulți, iar morfologia complexelor-K prezintă variații cu vârsta. Prin urmare, detectoarele de caracteristici internalizate de un model antrenat pe populații adulte nu sunt direct transferabile la date pediatrice, ceea ce produce erori în clasificare în special pentru stadiile N1, N2 și N3.

Arhitectura somnului se modifică odată cu vârsta. Proporțiile stadiilor de somn variază semnificativ de-a lungul vieții [3]. La nou-născuți, somnul REM poate reprezenta aproximativ 50% din totalul somnului, față de 20–25% la adulți, declinul producându-se treptat în copilărie. Somnul cu unde lente (N3) este mai abundent și mai intens la copiii mici și se reduce considerabil în adolescență. Aceste variații în distribuția claselor creează dezechilibre diferite de cele prezente în datele adulte, pe care modelele pre-antrenate nu sunt calibrate să le gestioneze.

Lipsa seturilor de date pediatrice de dimensiuni mari. Baze de date publice de adulți de referință HMC [1], cu 151 de pacienți, sau Sleep-EDF [7], cu 197 de înregistrări, nu au echivalent pediatric. Colectarea PSG la copii implică obstacole logistice și etice considerabile: consimțământul parental, toleranța scăzută față de echipamentul de monitorizare și nevoile specifice de supraveghere clinică. Consecința practică este că studiile de clasificare automată a somnului pediatric disponibile în literatură operează de regulă pe câteva zeci de pacienți o dimensiune insuficientă pentru antrenarea de novo a arhitecturilor cu milioane de parametri.

Absența comparațiilor sistematice pe pacienți pediatrici. Deși există lucrări izolate care abordează somnul pediatric, comparațiile directe între arhitecturi multiple, evaluate pe aceeași populație cu protocoale uniforme, nu au fost publicate în mod sistematic. Fără astfel de repere, nu este clar care arhitectură se comportă cel mai bine la populația pediatrică și în ce condiții. Lucrarea de față adresează în mod explicit această lacună.

1.3 Obiective

Lucrarea urmărește cinci obiective concrete:

(1) **Analiza și implementarea arhitecturilor de referință** DeepSleepNet [25] și AttnSleep [6], cu identificarea mecanismelor cheie, a aspectelor pozitive și a limitărilor

fiecăreia în contextul clasificării somnului. Pe lângă aceste modele de referință, sunt propuse **PediaSleepNet Base** și **PediaSleepNet V2**, două arhitecturi originale concepute în cadrul acestei lucrări, care integrează filtrare adaptivă depthwise pe grupe de vârstă, condiționare prin *Feature-wise Linear Modulation* (FiLM) [21] și, în versiunea V2, un cap auxiliar dedicat detecției stadiului N1.

(2) Pre-antrenarea celor patru arhitecturi pe seturi de date publice de adulți Sleep-EDF [7] și HMC [1] urmând specificațiile originale ale fiecărei lucrări, astfel încât evaluările ulterioare să fie comparabile cu rezultatele raportate în literatura de specialitate.

(3) Adaptarea modelelor prin *fine-tuning* pe un set de date pediatric, cu ajutorul tehnicilor de *transfer learning*. Această abordare permite valorificarea reprezentărilor dobândite pe populații adulte, compensând volumul redus de date disponibile la copii și adolescenți. Pentru PediaSleepNet Base și V2, etapa de *adaptare* urmează un protocol de dezghețare progresivă în trei etape, activând specific filtrele adaptive pe grupă de vârstă și modulele de condiționare FiLM.

(4) Compararea cantitativă a celor patru modele după *adaptarea* pe date pediatrice, utilizând acuratețe globală, *macro F1-score* și coeficientul Cohen's κ , completate de analiza erorilor per clasă prin matricea de confuzie.

(5) Evaluarea aportului mecanismelor de condiționare pe vârstă introduse în PediaSleepNet, filtre adaptive depthwise, modulele SE-FiLM și capul auxiliar N1 din V2, față de arhitecturi fără astfel de mecanisme, pentru a arăta câștigul adus de adaptarea explicită la semnalele EEG pediatrice.

Capitolul 2

Baze teoretice

2.1 Biologia somnului

Somnul este prezent la toate mamiferele studiate și sub forme asemănătoare la reptile și păsări, ceea ce sugerează o origine comună și funcții biologice cu adevărat esențiale [3]. Din perspectivă neurofiziologică, somnul nu reprezintă o reducere a activității cerebrale, ci o stare în care rețelele neuronale se reorganizează activ sub coordonarea unor sisteme subcorticale bine delimitate: nucleul suprachiasmatic principal oscilator circadian, aria preoptică ventrală și formațiunea reticulată ascendentă.

2.1.1 Structura temporală a somnului normal

La un adult sănătos, somnul nocturn se desfășoară în cicluri cu durată medie de aproximativ 90 de minute, repetate de 4–5 ori pe noapte. Fiecare ciclu începe cu parcurgerea treptată a stadiilor NREM (N1, N2, N3), urmată de un episod REM. Pe parcursul nopții, proporția somnului cu unde lente (N3) scade progresiv, în timp ce episoadele REM devin mai lungi, astfel că ultimele cicluri sunt dominate de somn REM [3].

Desfășurarea este diferită la **copii**: stadiul N3 este mai abundent și mai intens, cu unde delta de amplitudine ridicată; somnul REM reprezintă o fracție mai mare din total; ciclul de somn este mai scurt (circa 60 de minute la sugari, ajungând progresiv la 90 de minute în adolescență); și, la trezire, copiii mici prezintă o inerție de somn mai accentuată [3]. Aceste diferențe reflectă procesele active de maturizare cerebrală sinaptogenă, mielinizare, prunings sinaptic care se desfășoară de-a lungul copilăriei și se oglindesc direct în morfologia semnalului EEG.

2.2 Stadiile de somn și EEG

2.2.1 Principiile înregistrării EEG

Electroencefalografia (EEG) înregistrează câmpurile electrice generate de activitatea sincronă a populațiilor de neuroni corticali. Sursa principală a semnalului de pe scalp o reprezintă potențialele postsinaptice (PSP) ale neuronilor piramidali din straturile III și V ale cortexului, orientați perpendicular pe suprafața corticală [3]. Potențialele de acțiune individuale contribuie neglijabil la semnalul EEG de scalp, deoarece amplitudinea și durata lor sunt prea reduse pentru a genera un câmp detectabil la această distanță. Ceea ce se înregistrează este, prin urmare, suma spațio-temporală a mii de PSP produse simultan de populații neuronale largi.

Electrodul de pe scalp măsoară diferența de potențial față de un electrod de referință plasat uzual la nivelul mastoidelor. Amplitudinile tipice ale semnalului EEG sunt de ordinul zecilor de microvolți (μV), utilizând amplificatoare diferențiale cu câștig de 10^4 – 10^5 și cu un raport de respingere a modului comun (CMRR) > 100 dB pentru a separa semnalul util de zgomotele electrice ambientale. Pozițiile electrozilor sunt standardizate prin sistemul internațional 10–20, care utilizează repere anatomice craniene reproductibile; notația literă + număr identifică simultan regiunea anatomică și emisfera (numere impare pentru electrozi din emisfera stângă, pare pentru dreapta): Fp (frontopolar), F (frontal), C (central), P (parietal), O (occipital), T (temporal). Derivațiile Fpz-Cz și Pz-Oz, folosite în setul Sleep-EDF [15], sunt ambele bipolare de linie mediană și surprind activitate centrală și parieto-occipitală relevantă pentru toate stadiile de somn.

2.2.2 Benzile de frecvență EEG relevante pentru somn

Semnalul EEG este analizat pe benzi de frecvență cu semnificații fiziologice distincte [3, 12]:

- **Delta** (< 4 Hz): unde lente de amplitudine mare (75–200 μV), generate prin mecanisme talamo-corticale; caracterizează stadiul N3 (somnul profund). La copii, amplitudinile pot depăși 200 μV .
- **Theta** (4–8 Hz): activitate de amplitudine moderată, caracteristică stadiului N1, somnolenței și somnului REM; vizibilă mai ales la electrozii de vertex (Cz).
- **Alpha** (8–13 Hz): ritm de repaus, prezent occipital în starea de veghe relaxată cu ochii închiși; se atenuează la deschiderea ochilor sau la efort cognitiv.

- **Sigma** (12–15 Hz): cuprinde fuserile de somn (*sleep spindles*), semnătura stadiului N2. Sunt rafale sinusoidale cu durată 0,5–3 s, generate de circuitul talamo-cortical prin nucleul reticular talamic.
- **Beta** (> 15 Hz): activitate desincronizată de amplitudine mică, asociată cu starea de veghe activă și efortul cognitiv.
- **Gamma** (> 30 Hz): amplitudine foarte mică, asociată cu procesarea senzorială de înaltă frecvență; mai rar folosită în analiza somnului.

2.2.3 Stadiile de somn conform manualului AASM

Clasificarea clinică a somnului se bazează pe manualul publicat de American Academy of Sleep Medicine (AASM) [12], care definește cinci stadii distincte cu criterii multimodale EEG, EOG și EMG. Fiecare epocă (epoch) de 30 de secunde primește o singură etichetă de stadiu, iar clasificarea se realizează vizual de un specialist acreditat.

Stadiul W (Veghe). Semnătura EEG diferă în funcție de starea de alertă a pacientului: ritmul alfa (8–13 Hz), dominant occipital, apare atunci când pacientul este relaxat cu ochii închiși, în timp ce ochii deschiși produc activitate beta desincronizată de amplitudine mică [12]. Tonusul muscular EMG al bărbiei este ridicat, iar mișcările voluntare ale ochilor pot fi frecvente. La copii, frecvența de fond alfa poate fi cu 1–2 Hz mai joasă față de adulți, un efect al maturizării corticale incomplete [3].

Stadiul N1. Stadiul N1 marchează granița dintre veghe și somn și este cel mai dificil de clasificat cu acuratețe ridicată. Activitatea alfa scade sub 50% din durată epocii, lăsând loc activității theta de amplitudine mică în banda 2–7 Hz [12]. Mișcările lente ale globilor oculari (SEM) sunt tipice, iar undele ascuțite de vertex (*vertex sharp waves*), tranzienți cu amplitudini de 70–150 μ V, maximi central, pot apărea sporadic. Incertitudinea spectrală față de stadiile W și N2 explică ratele de eroare sistematic mai ridicate pe care clasificatoarele automate le înregistrează pentru N1.

Stadiul N2. Somnul devine stabil la N2. Criteriul minim pentru atribuirea acestui stadiu este prezența cel puțin a unui complex-K sau a unui fus de somn pe parcursul epocii [12]. **Complexele-K** sunt deflexiuni bifazice negativ-pozitive cu amplitudine $\geq 75 \mu$ V și durată $\geq 0,5$ s, cu distribuție maximă la electrozii frontali; ele apar fie spontan, fie ca răspuns la stimuli externi bruști. **Fuserile de somn** sunt rafale oscilante în banda 12–15 Hz la adulți (13–15 Hz la copii), cu durată 0,5–3 s și distribuție maximă centroparietal [12, 3].

Stadiul N3 (Slow Wave Sleep, SWS). Conform criteriilor AASM [12], stadiul N3 este definit prin prezența undelor lente cu frecvența 0,5–2 Hz și amplitudine $\geq 75 \mu$ V în

cel puțin 20% din durata epocii. La copii, activitatea delta este natural mai amplă decât la adulți, unde pot atinge 300–500 μV , ceea ce face ca pragul de 75 μV conceput pentru adulți să fie atins mai ușor, cu riscul de a supraestima procentul de N3 la vârste mici. Aceasta este una dintre diferențele care complică transferul direct al modelelor antrenate pe date de adulți.

Stadiul REM (Rapid Eye Movement). EEG-ul din REM prezintă activitate de amplitudine mică cu frecvențe mixte, unde alfa de frecvență ușor mai joasă față de veghe și undele caracteristice “dinte de fierăstrău” (2–6 Hz), vizibile mai ales frontal [12]. EOG înregistrează mișcări rapide și conjugate ale ochilor, iar EMG al bărbiei atinge cel mai scăzut nivel din întreaga noapte, atonia musculară este generată prin inhibiție pontină a motoneuronilor spinali, ceea ce face din REM stadiul cel mai ușor de identificat prin semnal EMG [3]. La copii, perioadele REM sunt distribuite mai uniform pe parcursul nopții față de adulți.

2.2.4 Particularitățile EEG pediatric

Maturizarea cortexului cerebral modifică semnificativ profilul EEG pe parcursul copilăriei și adolescenței, cu implicații directe asupra performanței modelelor automate [3]:

Frecvența de fond crește treptat de la circa 6–7 Hz la sugari la 8–9 Hz la copii de 3 ani, atingând valorile adulte de 10–12 Hz abia în adolescența târzie. Markerii clasici ai stadiului W bazați pe alfa la 10 Hz sunt astfel deplasați spectral la vârste mici.

Amplitudinile EEG sunt în general mai mari la copii, în special undele delta din N3, care pot atinge 300–500 μV față de pragul de 75 μV definit de AASM pentru adulți.

Fusurile de somn sunt mai frecvente per epocă la copii față de adulți, au o frecvență centrală ușor mai ridicată și o distribuție mai anterioară pe scalp. Morfologia complexelor-K prezintă de asemenea variații cu vârsta.

Somnul REM scade de la ~50% la nou-născuți la ~20–25% în adolescență, ajungând la valorile adulte. Toate aceste diferențe justifică necesitatea adaptării modelelor prin *adaptarea* pe date pediatrice, abordare centrală a lucrării de față.

Capitolul 3

State of the Art

Metodele tradiționale de clasificare automată a stadiilor de somn se bazau pe extracția manuală de trăsături din domeniul timp, frecvență sau timp-frecvență (STFT, transformata wavelet), urmate de clasificatoare clasice, SVM, Random Forest sau rețele neuronale puțin adânci. Deși aceste abordări au produs rezultate rezonabile pe experimente controlate, performanța lor scade considerabil atunci când sunt aplicate pe date din spitale sau echipamente diferite de cele pe care au fost antrenate. Rețelele de deep learning au schimbat această dinamică: prin antrenarea direct pe semnalul EEG brut, ele pot dobândi reprezentări adaptate datelor fără intervenție umană în selecția trăsăturilor [25, 6].

Capitolul de față descrie cele patru arhitecturi care formează esența metodologică a lucrării: **DeepSleepNet** [25], **AttnSleep** [6], **PediaSleepNet Base** și **PediaSleepNet V2** (arhitecturi proprii). Fiecare reprezintă un răspuns la limitările predecesorilor, introducând inovații arhitecturale orientate spre problemele specifice clasificării somnului.

3.1 DeepSleepNet

DeepSleepNet, modelul propus de Supratak et al. [25], a reprezentat la momentul publicării un pas metodologic important: o arhitectură end-to-end capabilă să clasifice stadiile de somn direct din semnalul EEG brut monocanal, eliminând etapa de extracție manuală a trăsăturilor. Mai mult, problema este formulată ca etichetare de secvențe, nu ca o clasificare de tip independentă epocă cu epocă, permițând modelului să internalizeze implicit constrângeri de tranziție între stadii (de exemplu, că N3 nu urmează direct W).

3.1.1 Arhitectura modelului

DeepSleepNet este organizat în două module care acoperă funcții complementare **Representation Learning** pentru extracția de trăsături și **Sequence Residual Learning** pentru modelarea contextului temporal (Figura 3.1). Intrarea este o epocă EEG brută de 30 de secunde $\mathbf{x}_i$, iar ieșirea este distribuția de probabilitate peste cele cinci stadii AASM: W, N1, N2, N3 și REM.

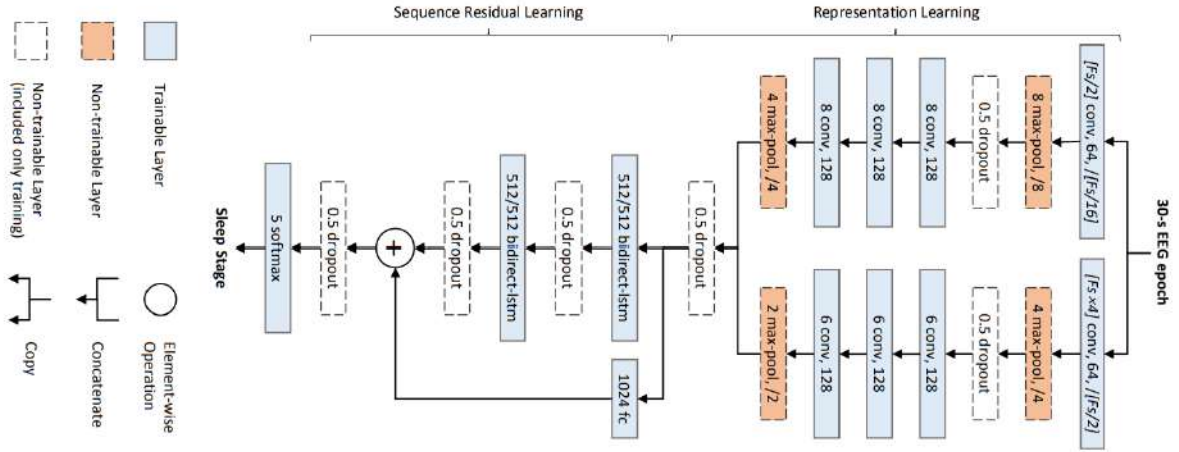


Figura 3.1 Arhitectura DeepSleepNet. Modulul de Representation Learning conține două rețele CNN paralele cu filtre de dimensiuni diferite, care extrag trăsături temporale și frecvențiale din epoca EEG brută. Modulul Sequence Residual Learning utilizează LSTM bidirecționale cu conexiune reziduală pentru a modela regulile de tranziție dintre stadii. Adaptat după [25].

Modulul de Representation Learning

Extracția trăsăturilor se realizează prin **două rețele convoluționale paralele** (CNN-1 și CNN-2) cu dimensiuni diferite ale filtrelor la primul strat, o alegere de design motivată de nevoia de a acoperi simultan componente EEG de scări temporale foarte diferite [25]. CNN-1 folosește filtre de dimensiunea $F_s/2$ (fereastră scurtă de 0,5 s), optimizate pentru a detecta evenimente locale precum complexe-K și fusurile de somn. CNN-2 folosește filtre de dimensiunea $F_s \times 4$ (fereastră de 4 s), mai potrivite pentru componentele delta lente din stadiul N3. Fiecare ramură include patru straturi convoluționale, două de max-pooling, *batch normalization* [13] și activare ReLU [19]. Reprezentările rezultate sunt concatenate astfel:

$$\mathbf{a}_i = \text{CNN}_{\theta_s}(\mathbf{x}_i) \parallel \text{CNN}_{\theta_l}(\mathbf{x}_i) \in \mathbb{R}^{1024}, \quad (3.1)$$

unde $\mathbf{x}_i$ este epoca EEG brută de la intrare, CNN_{θ_s} și CNN_{θ_l} sunt cele două ramuri convoluționale paralele cu filtre scurte (θ_s) respectiv lungi (θ_l), $\parallel$ denotă operația de concatenare pe axa trăsăturilor, iar $\mathbf{a}_i \in \mathbb{R}^{1024}$ este vectorul de trăsături rezultat pentru epoca i .

Modulul de Sequence Residual Learning

Contextul temporal, constrângerile de tranziție dintre stadii, este capturat prin **LSTM bidirecționale** [10, 23] cu conexiuni peephole, la care se adaugă o **conexiune reziduală**

[8]:

$$\mathbf{o}_t = \mathbf{h}_t^f \parallel \mathbf{h}_t^b + \text{FC}_\theta(\mathbf{a}_t), \quad (3.2)$$

unde $\mathbf{h}_t^f$ și $\mathbf{h}_t^b$ sunt stările ascunse ale LSTM-ului în direcția înainte (*forward*) respectiv înapoi (*backward*) la pasul de timp t , $\parallel$ denotă concatenarea, iar $\text{FC}_\theta(\mathbf{a}_t)$ este conexiunea reziduală care proiectează liniar trăsăturile $\mathbf{a}_t$ extrase de modulul de *representation learning*. Vectorul $\mathbf{o}_t$ este transmis unui strat softmax cu cinci ieșiri, producând predicția $\hat{y}_t$ pentru fiecare epocă.

3.1.2 Limitări

Numărul mare de parametri ($\sim 22,6$ milioane) face ca DeepSleepNet să necesite volume mari de date de antrenare; performanța scade semnificativ pe seturi mici. Filtrele CNN optimizate pentru un canal EEG specific generalizează slab pe canale sau populații diferite fără reantrenare. De asemenea, procesarea secvențială bidirecțională presupune disponibilitatea întregii secvențe de L epoci înainte de predicție, ceea ce îl face nepotrivit pentru aplicații cu constrângeri de latență.

3.2 AttnSleep

AttnSleep, propus de Eldele et al. [6], înlocuiește componentele recurente cu mecanisme de atenție și convoluții cauzale, permițând procesare paralelă a trăsăturilor. Modelul introduce totodată un mecanism de recalibrare adaptivă a trăsăturilor și o funcție de pierdere cu ponderi de clasă logaritmice pentru a compensa dezechilibrul de clase fără augmentare sintetică.

3.2.1 Arhitectura modelului

AttnSleep procesează o singură epocă EEG de 30 de secunde și produce distribuția de probabilitate pe cele cinci stadii. Arhitectura, prezentată în Figura 3.2, conține trei blocuri principale: **Multi-Resolution CNN** (MRCNN), **Adaptive Feature Recalibration** (AFR) și **Temporal Context Encoder** (TCE).

Multi-Resolution CNN (MRCNN)

MRCNN conține două ramuri CNN paralele cu nuclee de dimensiuni diferite [6]. Pentru $F_s = 100$ Hz: **ramura cu nucleu mic** (dimensiune 50, fereastră 0,5 s) vizează benzile alfa și theta (4–12 Hz), iar **ramura cu nucleu mare** (dimensiune 400, fereastră 4 s) vizează

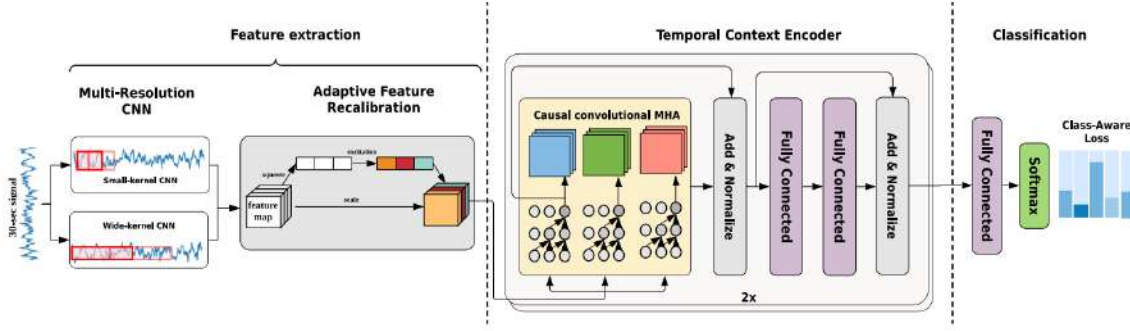


Figura 3.2 Arhitectura AttnSleep. MRCNN extrage trăsături prin două ramuri CNN cu nuclee de dimensiuni diferite, captând informații din benzi de frecvență EEG distincte. AFR recalibrează inter-dependențele dintre trăsături printr-un bloc *Squeeze-and-Excitation* rezidual. TCE captează dependențele temporale prin atenție multi-head cu convoluții cauzale, fără componente recurente. Adaptat după [6].

banda delta (<4 Hz), caracteristică stadiului N3. Fiecare ramură conține straturi Conv1D cu *batch normalization* [13], activare GELU [9] și max-pooling. Ieșirile sunt concatenate în harta de trăsături $\mathbf{I} \in \mathbb{R}^{N \times d}$.

Adaptive Feature Recalibration (AFR)

AFR modelează inter-dependențele dintre trăsăturile extrase de MRCNN printr-un bloc **Squeeze-and-Excitation rezidual** [11]. Operațiile de comprimare globală (*squeeze*) și estimare a importanței relative (*excitație*) sunt:

$$\mathbf{e} = \sigma(\mathbf{W}_2 \cdot \delta(\mathbf{W}_1 \cdot \mathbf{s})), \quad \mathbf{X} = \mathbf{I} + \mathbf{F} \otimes \mathbf{e}, \quad (3.3)$$

unde $\mathbf{s}$ este reprezentarea global obținută prin comprimarea (*squeeze*) hărții de trăsături de intrare $\mathbf{I}$, $\mathbf{W}_1$ și $\mathbf{W}_2$ sunt matricele de ponderi ale celor două straturi complet conectate ale blocului de excitație, δ este activarea ReLU, σ este funcția sigmoidă, $\mathbf{e}$ este vectorul de ponderi de recalibrare per canal, $\mathbf{F}$ este harta de trăsături supusă scalării, $\otimes$ denotă înmulțirea pe canale (*channel-wise*), iar $\mathbf{X}$ este harta de trăsături recalibrată. Conexiunea reziduală adaugă trăsăturile originale nemodificate, prevenind pierderea de informație prin recalibrare.

Temporal Context Encoder (TCE)

TCE înlocuiește componentele recurente cu **Multi-Head Attention** (MHA) [26] și convoluții cauzale:

$$\text{MHA}(\mathbf{X}, \mathbf{X}, \mathbf{X}) = \text{Concat}(A^1, \dots, A^H), \quad A^h = \text{ATT}(\hat{X}^h, \hat{X}^h, \hat{X}^h), \quad (3.4)$$

unde $\mathbf{X}$ este harta de trăsături de la intrare, folosită simultan ca interogare (*query*), cheie (*key*) și valoare (*value*), H este numărul de capete de atenție, A^h este ieșirea capului h , $\hat{X}^h$ sunt proiecțiile liniare corespunzătoare capului h , $\text{ATT}(\cdot)$ este atenția de tip produs scalar scalat (*scaled dot-product attention*), iar Concat concatenează ieșirile tuturor celor H capete. TCE pune două astfel de structuri cu normalizare layer [2] și conexiuni reziduale. Față de LSTM, convoluțiile cauzale reduc semnificativ timpul de antrenare: circa 21 de minute per fold față de 2,5 ore pentru DeepSleepNet pe Sleep-EDF-20 [25, 6].

Funcția de pierdere cu ponderi logaritmice

Pentru tratarea dezechilibrului de clase, AttnSleep propune o funcție de pierdere cu ponderi de clasă logaritmice [6]:

$$\mathcal{L} = -\frac{1}{M} \sum_{k=1}^K \sum_{i=1}^M w_k y_i^k \log(\hat{y}_i^k), \quad w_k = \mu_k \cdot \max\left(1, \log\left(\frac{\mu_k M}{M_k}\right)\right), \quad (3.5)$$

unde M este numărul total de epoci din lotul de antrenare, $K = 5$ este numărul de clase, y_i^k este eticheta reală a epocii i pentru clasa k , $\hat{y}_i^k$ este probabilitatea prezisă de model pentru aceeași epocă și clasă, w_k este ponderea logaritmică asociată clasei k , μ_k este un factor de scalare, iar M_k este numărul de epoci aparținând clasei k ; raportul $\mu_k M / M_k$ crește ponderea claselor sub-reprezentate, compensând astfel dezechilibrul de clase.

3.2.2 Limitări

În versiunea single-epoch, AttnSleep procesează fiecare epocă independent, fără modelarea explicită a contextului secvențial, o limitare față de modele sequence-to-sequence precum DeepSleepNet [25]. Versiunea multi-epoch (3 epoci) ameliorează parțial această problemă. Mecanismul de atenție poate fi dominat de pozițiile cu amplitudini mari, pierzând tipare EEG locale de amplitudine redusă.

3.3 PediaSleepNet (modele proprii)

În cadrul acestei lucrări sunt propuse două versiuni ale unei noi arhitecturi originale, denumite **PediaSleepNet Base** și **PediaSleepNet V2**, concepute special pentru clasificarea automată a stadiilor de somn la copii și adolescenți pornind de la un singur canal EEG brut. Față de DeepSleepNet [25] și AttnSleep [6], arhitectura PediaSleepNet introduce mecanisme dedicate de adaptare la specificul neurofiziologic și spectral al semnalelor EEG din populația pediatrică.

Ambele modele prelucrează epoci EEG brute de 30 de secunde, eșantionate la $F_s = 256$ Hz (7680 de eșantioane per epocă), și operează cu două grupe demografice: **G0** (copii sub 13 ani) și **G1** (adolescenți și tineri de 13 ani și peste).

3.3.1 PediaSleepNet Base

Modelul de bază este structurat modular, integrând straturi de filtrare adaptivă condiționată, extracție multi-rezoluție și un codicator al contextului temporal global. Arhitectura acestui model este prezentată în Figura 3.3.

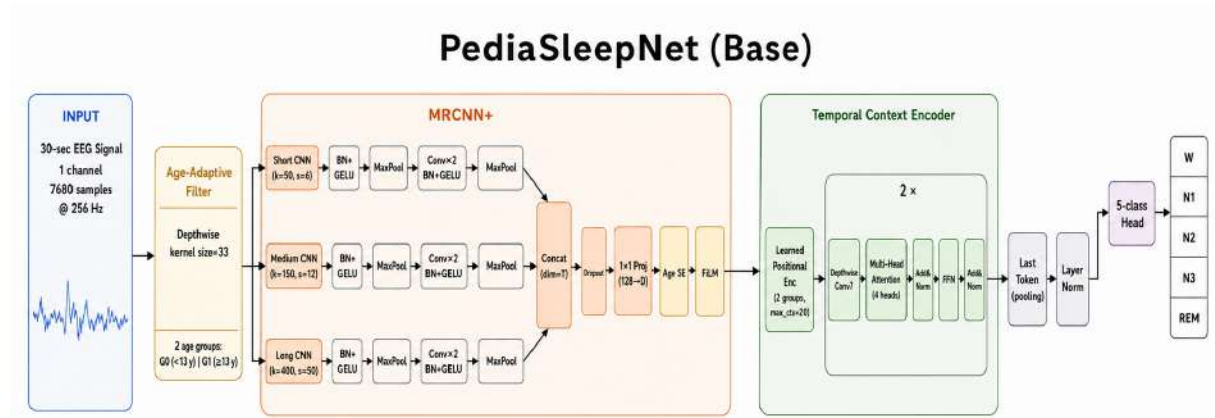


Figura 3.3 Arhitectura detaliată a modelului PediaSleepNet Base, evidențiind filtrarea adaptivă pe grupe de vârstă, ramurile MRCNN+ concatenate temporal și codicatorul TCE. Sursa: elaborare proprie.

Filtrarea adaptivă în funcție de vârstă (Age-Adaptive Depthwise Filter)

Înainte ca semnalul să ajungă la blocurile convoluționale principale, acesta este trecut printr-un strat de filtrare convoluțională de tip *depthwise* [4] cu seturi complet separate de parametri pentru grupele G0 și G1. Dimensiunea nucleului este $k = 33$ (≈ 130 ms la 256 Hz), acoperind cicluri complete ale oscilațiilor din banda sigma (fusuri de somn, 12–15 Hz) și permițând filtrului să funcționeze ca un filtru trece-bandă dinamic adaptat spectrului EEG caracteristic fiecărei grupe de vârstă. Activarea utilizată după acest strat este GELU [9].

Extracția de trăsături la rezoluții multiple (MRCNN+)

Semnalul filtrat adaptiv este trimis în paralel către modulul MRCNN+, configurat pe trei ramuri convoluționale independente cu nuclee de dimensiuni diferite:

- **Ramura cu nucleu mic** ($k = 50$, pas 6): specializată în capturarea fenomenelor tranzitorii rapide și a micro-structurilor somnului (fusuri de somn și activitate alfa/theta rapidă).
- **Ramura cu nucleu mediu** ($k = 150$, pas 12): echilibrează rezoluția temporală cu cea spectrală pentru analiza oscilațiilor intermediare theta–delta.
- **Ramura cu nucleu mare** ($k = 400$, pas 50): optimizată pentru captarea macro-structurilor și a undelor delta lente (< 4 Hz), fundamentale pentru identificarea stadiului N3.

Fiecare ramură conține: Conv1D $\rightarrow$ BatchNorm $\rightarrow$ GELU $\rightarrow$ MaxPool $\rightarrow$ Dropout(0,25), urmată de două straturi Conv1D–BN–GELU și un MaxPool final. Hărțile de trăsături ale celor trei ramuri sunt redimensionate prin interpolare la o dimensiune temporală comună și **concatenate pe axa temporală**, după care sunt proiectate la $d_{\text{model}} = 128$ printr-un strat Conv1D cu nucleu 1×1 .

Condiționarea FiLM și Squeeze-and-Excitation

Reprezentarea proiectată trece printr-un modul *Squeeze-and-Excitation* (SE) [11] pentru recalibrarea dinamică a importanței canalelor de trăsături, vectorul de pooling fiind condiționat prin concatenare cu un embedding de vârstă de dimensiune 64. Ulterior, un strat FiLM (*Feature-wise Linear Modulation*) [21] aplică modularea finală prin parametri de scalare (γ) și deplasare (β) învățați per grupă:

$$\text{FiLM}(\mathbf{x}, g) = (1 + \gamma_g) \odot \mathbf{x} + \beta_g, \quad (3.6)$$

unde $\mathbf{x}$ este reprezentarea de trăsături de la intrare, $g \in \{G0, G1\}$ este grupa de vârstă, γ_g și β_g sunt parametrii de scalare respectiv de deplasare învățați separat pentru fiecare grupă, iar $\odot$ denotă înmulțirea element cu element. Formularea $(1 + \gamma_g)$ asigură că la inițializare ($\gamma_g = 0$) modulul aplică transformarea identitate, lăsând trăsăturile nemodificate.

Codificatorul Temporal (Temporal Context Encoder – TCE)

Dependențele pe termen lung dintre epocile consecutive sunt modelate de TCE, compus din $N = 2$ straturi Transformer cu $h = 4$ capete de atenție [26]. Hărțile de trăsături intra-epocă (inter-epoch) sunt mai întâi reduse la un singur vector prin *mean pooling*, iar TCE aplică self-attention între embedding-urile epocilor de-a lungul axei de secvență $(B, \text{seq_len}, D)$, unde $\text{seq_len} = 20$ epoci consecutive. Înaintea calculului de atenție se aplică o convoluție locală depthwise ($k = 7$) pentru a păstra logica secvențială locală a

tranzițiilor dintre stadii. Clasificarea finală a stadiilor (W, N1, N2, N3, REM) se realizează la nivelul ultimului token normalizat, printr-un cap liniar cu 5 ieșiri.

3.3.2 PediaSleepNet V2

PediaSleepNet V2 a fost gândit ca o evoluție arhitecturală a versiunii Base, cu obiective de optimizare a eficienței computaționale și de creștere a performanței de detecție pentru stadiul N1. Arhitectura completă, cu modificările față de versiunea Base marcate explicit, este prezentată în Figura 3.4.

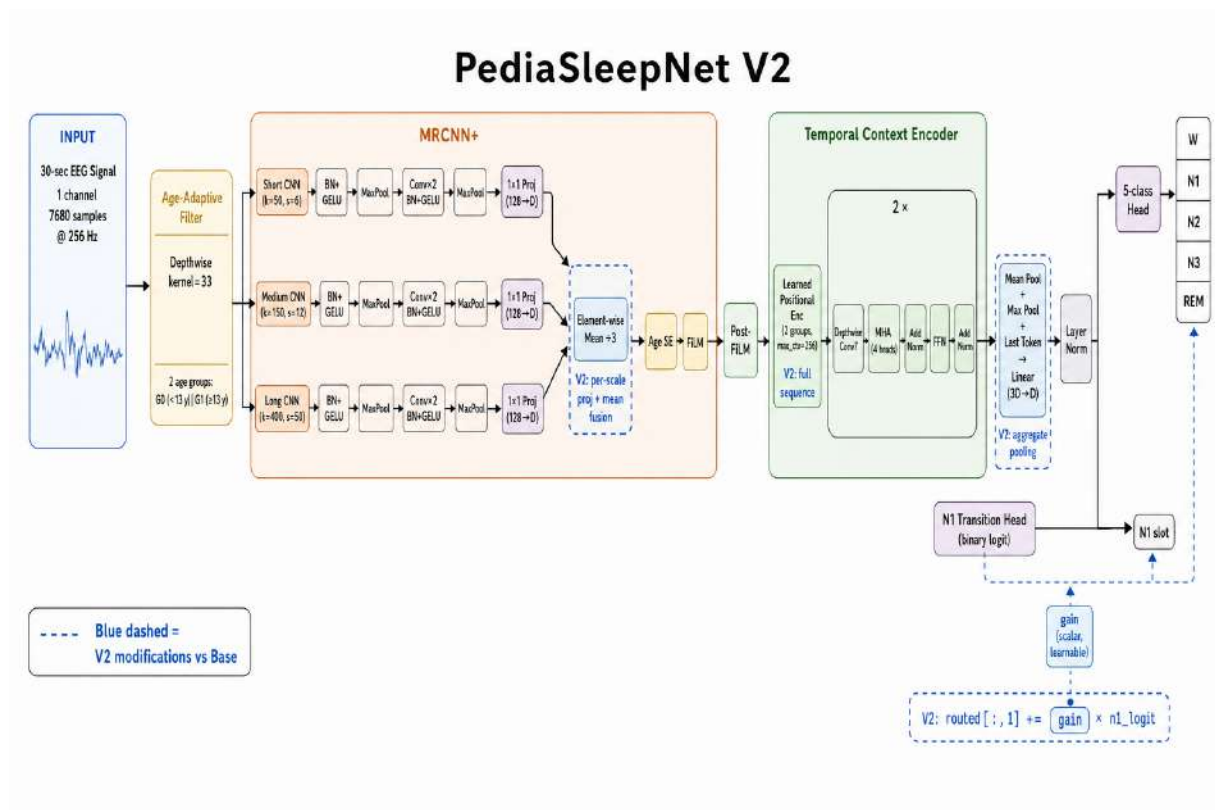


Figura 3.4 Arhitectura modelului PediaSleepNet V2. Modificările față de versiunea Base (marcate cu linii punctate albastre) sunt: proiecția liniară individuală per ramură MRCNN+, fuziunea prin medie element-wise și capul auxiliar de detecție N1. Sursa: elaborare proprie.

V2 păstrează neschimbate formatele de intrare, gruparea demografică ($G0$, $G1$), filtrarea adaptivă și structura TCE, însă introduce trei modificări structurale majore:

Proiectare individuală în modulul MRCNN+

Spre deosebire de versiunea Base, unde ramurile sunt concatenate înainte de o proiecție comună, în V2 fiecare dintre cele trei ramuri MRCNN+ este proiectată individual la $d_{\text{model}} = 128$ printr-un strat Conv1D 1×1 propriu, înainte de fuziune.

Fuziune prin medie element-wise (Element-wise Mean Fusion)

În locul concatenării temporale din versiunea Base, care tripla artificial lungimea secvenței transmise TCE, V2 combină cele trei reprezentări proiectate prin medie element cu element:

$$\mathbf{X}_{\text{fused}} = \frac{1}{3} \sum_{k \in \{50, 150, 400\}} \text{Conv1D}_{1 \times 1}^{(k)}(\mathbf{X}_k), \quad (3.7)$$

unde $\mathbf{X}_k$ este harta de trăsături produsă de ramura MRCNN+ cu nucleul de dimensiune $k \in \{50, 150, 400\}$, $\text{Conv1D}_{1 \times 1}^{(k)}$ este proiecția liniară individuală a acelei ramuri la $d_{\text{model}} = 128$, iar $\mathbf{X}_{\text{fused}}$ este reprezentarea concatenată obținută prin media element cu element a celor trei ramuri. Această abordare reduce dimensionalitatea operanzilor din straturile de atenție ale TCE și constrânge rețeaua să integreze trăsăturile complementare la nivelul fiecărui pas temporal, menținând modelul sub 1 M de parametri.

Capul auxiliar de detecție N1

Stadiul N1 suferă de reprezentare cantitativă redusă în înregistrările PSG și de ambiguitate vizuală ridicată pe EEG. V2 introduce un cap de clasificare auxiliar dedicat detecției binare a stadiului N1 (*N1 Transition Head*), antrenat printr-o funcție de pierdere secundară Binary Cross-Entropy. Logitul binar generat este injectat direct în slotul clasei N1 din capul principal de clasificare, printr-un coeficient scalar α învățat automat:

$$\tilde{l}_{\text{N1}} = l_{\text{N1}} + \alpha \cdot l_{\text{aux, N1}}, \quad (3.8)$$

unde l_{N1} este logitul clasei N1 produs de capul principal de clasificare, $l_{\text{aux, N1}}$ este logitul binar produs de capul auxiliar de detecție N1, α este un coeficient scalar învățat automat care controlează contribuția capului auxiliar, iar $\tilde{l}_{\text{N1}}$ este logitul N1 final injectat înapoi în slotul clasei N1 din capul principal. Această cuplare structurală forțează gradientii să propage informație discriminativă specifică stadiului N1 înapoi în rețea.

Principalii hiperparametri ai arhitecturilor PediaSleepNet sunt detaliați în Anexa A.1 (Tabelul A.1).

3.3.3 Limitări

Validarea modelelor a fost realizată pe un set de date pediatric local cu număr limitat de subiecți (19 pacienți, 9–17 ani), fiind necesară testare viitoare pe cohorte clinice mai extinse. Segmentarea filtrelor adaptive pe bucle condiționale în funcție de grupul demografic limitează parțial paralelizarea la nivel de GPU în timpul execuției batch-urilor mixte.

3.4 Sinteză comparativă

Cele patru arhitecturi analizate în această lucrare propun abordări tehnice fundamentale diferite pentru aceeași sarcină. DeepSleepNet [25] a validat abordarea end-to-end cu modelare secvențială bidirecțională BiLSTM, cu prețul unui număr masiv de parametri ($\sim 22,6$ M). AttnSleep [6] a demonstrat că mecanismele de atenție din Transformer pot înlocui componentele recurente cu o reducere dramatică a parametrilor (~ 523 K), procesând însă fiecare epocă independent. PediaSleepNet extinde aceste fundamente introducând adaptarea la vârstă: filtrul depthwise condițional și modulul SE-FiLM extrag caracteristici legate direct de neurofiziologia pediatrică. Versiunea V2 rafinează arhitectura prin fuziunea element-wise (model compact sub 1 M parametri) și capul auxiliar N1 cu rutare de gradienti, adresând direct principala slăbiciune a sistemelor automate de clasificare a somnului pediatric.

Proprietate	DeepSleepNet	AttnSleep	PSN Base	PSN V2
Framework	TensorFlow 1.x	PyTorch	PyTorch	PyTorch
Volum parametri	$\sim 22,6$ M	~ 523 K	$\sim 1,1$ M	~ 945 K
F_s nativă	100 Hz	100 Hz	256 Hz	256 Hz
Ramuri CNN	2 (Dual)	2 (Dual)	3 (Multi-Res)	3 (Multi-Res)
Fuziune trăsături	Concatenare	Atenție	Concatenare	Medie elem.-wise
Context temporal	BiLSTM	Transformer	Conv+Transformer	Conv+Transformer
Fereastră context	25 epoci	1 epocă	20 epoci	20 epoci
Condiționare vârstă	Nu	Nu	Da (SE+FiLM)	Da (SE+FiLM)
Optimizator N1	Nu	Nu	Nu	Da (cap binar+ α)

Tabel 3.1 Comparatie sintetică a arhitecturilor evaluate în lucrare.

Capitolul 4

Seturi de date

Experimentele din această lucrare se bazează pe trei seturi de date, alese pentru a acoperi atât etapa de pre-antrenare, cât și etapa de *adaptare* pe date pediatrice. **Sleep-EDF Expanded** [7, 15] și **HMC Sleep Staging** [1] conțin înregistrări de la adulți și sunt disponibile public, servind ca surse de pre-antrenare. Al treilea set conține înregistrări pediatrice de la copii cu vârste între 9 și 17 ani și constituie sursa pentru evaluarea transferului între domenii al modelelor **DeepSleepNet** [25], **AttnSleep** [6] și **PediaSleepNet**.

4.1 Sleep-EDF

Sleep-EDF Expanded [7, 15] este una dintre cele mai utilizate baze de date PSG publice, disponibilă pe PhysioNet, cu 197 de înregistrări de noapte întreagă. Cuprinde două sub-studii: **Sleep Cassette (SC)**, 153 de înregistrări de la persoane sănătoase (25–101 ani), colectate la domiciliu în 1987–1991; și **Sleep Telemetry (ST)**, 44 de înregistrări din 1994 privind efectele temazepamului versus placebo la 22 de pacienți.

Semnalele EEG (derivații **Fpz-Cz** și **Pz-Oz**), EOG și EMG sunt eșantionate la **100 Hz** [15]. Hipnogramele sunt clasificate conform standardului Rechtschaffen & Kales (R&K) [22], cu epoci de 30 de secunde, în format **EDF+** [14]. Subseturile cele mai frecvent utilizate în literatura de clasificare automată sunt **Sleep-EDF-20** (primii 20 de pacienți SC) și **Sleep-EDF-78** (78 de pacienți SC), adoptate și de lucrările de referință [25, 6] pentru comparabilitate inter-studii.

4.2 HMC Sleep Staging

Haaglanden Medisch Centrum (HMC) Sleep Staging conține 151 de înregistrări PSG de noapte întreagă colectate în 2018 la spitalul HMC din Haga, Olanda, publicat pe PhysioNet în 2022. Spre deosebire de Sleep-EDF, HMC include o populație heterogenă de pacienți cu tulburări de somn (85 bărbați, 66 femei, vârstă medie $53,9 \pm 15,4$ ani), mai reprezentativă pentru contextul clinic real.

Fiecare înregistrare conține patru derivații EEG (**F4/M1**, **C4/M1**, **O2/M1**, **C3/M2**), două derivații EOG, EMG și ECG, eșantionate la **256 Hz** și stocate în format EDF [14]. Hipnogramele sunt clasificate conform **AASM v2.4** [12] (cinci stadii: W, N1, N2, N3, R). Versiunea utilizată este **v1.1**, cu **148 de înregistrări** utilizabile (trei înregistrări eliminate din cauza erorilor de semnal raportate în setul original).

HMC completează Sleep-EDF prin adoptarea standardului AASM modern, rata de eșantionare mai ridicată (256 Hz față de 100 Hz) și montajul multi-electrod cu patru derivații EEG, oferind o perspectivă clinică mai diversă față de pacienții sănătoși din Sleep-EDF.

4.3 Setul de date pediatric

Setul de date pediatric constituie sursa de date pentru etapa de *adaptare*. Acesta conține **19 înregistrări PSG de noapte întreagă** de la copii și adolescenți cu vârste între 9 și 17 ani (medie $13,9 \pm 2,2$ ani), 10 băieți și 9 fete, cu durata medie de $8,9 \pm 0,5$ ore. Datele sunt stocate în format **EDF** [14], cu 27 de canale per înregistrare (EEG, EOG, EMG și alte semnale fiziologice). Amplitudinile EEG ating până la $\pm 5000 \mu V$, consistente cu amplitudinile intrinsec mai ridicate ale EEG pediatric [3].

Înregistrările provin de la **Spitalul Clinic de Copii „Victor Gomoiu”**, unde pacienții au fost trimiși pentru investigație polisomnografică din cauza **cefaleelor și migrenelor**. Pentru a reduce variabilitatea legată de condiții neurologice care alterează profilul EEG, au fost excluși copiii cu epilepsie, tulburări din spectrul autist, anomalii structurale cerebrale sau traumatism cranio-cerebral. Din perspectiva funcției respiratorii în timpul somnului, un singur pacient prezenta apnee, restul de 18 având o respirație normală pe parcursul înregistrării.

Stadiile de somn au fost adnotate de personal medical de la Spitalul Clinic de Copii „Victor Gomoiu” conform standardului **American Academy of Sleep Medicine (AASM)** [12]. Toate semnalele au fost achiziționate la o rată de eșantionare de **512 Hz**; nu s-au aplicat filtre suplimentare dincolo de filtrul trece-jos hardware, cu frecvența de tăiere la 256 Hz. În etapa de preprocesare, epocile sunt reeșantionate la rata țintă a fiecărui model (100 Hz sau 256 Hz), conform pipeline-ului descris în secțiunea 5.1.

Statisticile PSG sunt rezumate în Tabelul 4.1. Distribuția de vârstă (Figurile 4.1–4.2) acoperă intervalul 9–17 ani, cu o concentrare în jurul vârstei de 12–16 ani și fără diferențe semnificative de vârstă între sexe ($p = 0,69$).

Stadiul N2 este dominant (54,3% TST), iar stadiul N1 este cel mai redus (9,4%), dezechilibru similar cu cel din seturile adulte. Stadiul N3 prezintă variabilitate inter-

Parametru	Medie $\pm$ SD	[Min, Max]
Vârstă (ani)	13,9 $\pm$ 2,2	[9, 17]
Durată înregistrare (ore)	8,9 $\pm$ 0,5	[8.2, 10.2]
Eficiență somn (%)	87,5 $\pm$ 10,9	[50.8, 96.4]
Total arousals	25,1 $\pm$ 9,9	[11, 47]
Total treziri	10,8 $\pm$ 4,9	[3, 19]
Tranziții inter-stadii	64,4 $\pm$ 19,1	[38, 107]
N1 (% TST)	9,4 $\pm$ 4,1	[4, 21]
N2 (% TST)	54,3 $\pm$ 5,9	[45, 62]
N3 (% TST)	17,2 $\pm$ 5,3	[7, 30]
REM (% TST)	19,1 $\pm$ 3,1	[11, 24]

Tabel 4.1 Statistici descriptive ale parametrilor PSG pentru setul de date pediatric ($N = 19$). Valorile sunt prezentate ca medie $\pm$ abatere standard, cu intervalul [minim, maxim].

individuală ridicată (7–30%), reflectând diferențele individuale de maturizare a somnului cu unde lente în intervalul 9–17 ani. Fetele prezintă o proporție semnificativ mai mare de N3 față de băieți (20,1% vs. 14,8%, $p = 0,036$), aspect relevant pentru stratificarea seturilor de antrenare/test.

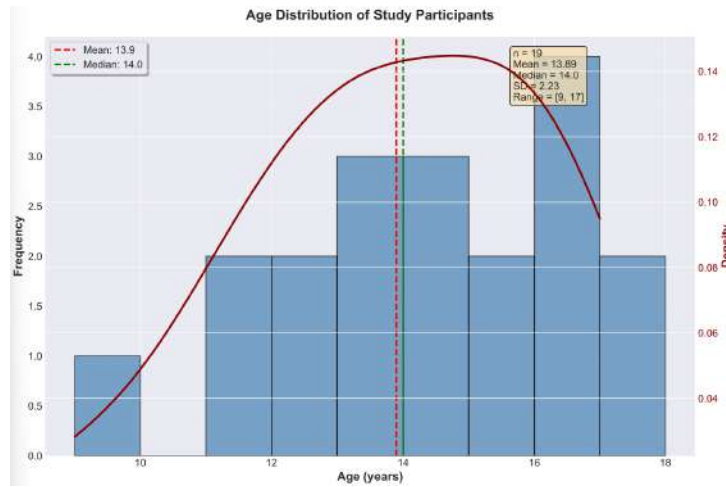


Figura 4.1 Distribuția vârstelor în setul de date pediatric ($N = 19$, vârste 9–17 ani). Histograma ilustrează concentrarea pacienților în intervalul 12–16 ani. Sursa: elaborare proprie.

Dimensiunea redusă a setului (19 pacienți) exclude antrenarea de novo a arhitecturilor de deep learning, justificând abordarea de *transfer learning* adoptată în această lucrare

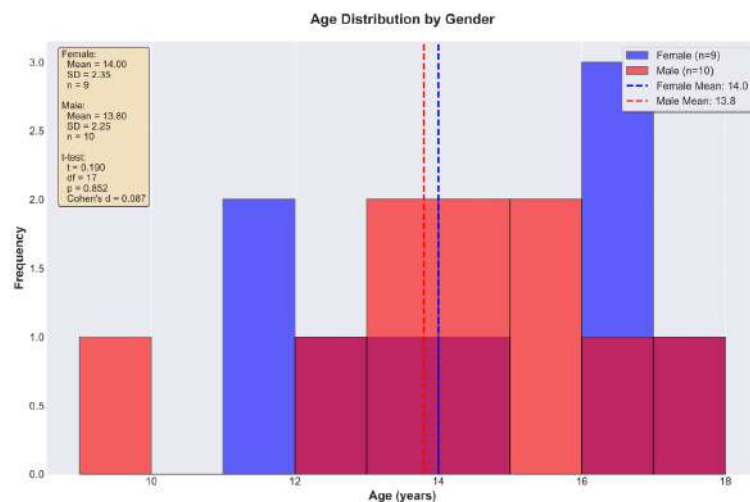


Figura 4.2 Distribuția vârstelor separat pe gen ($p = 0,69$, test t independent). Nu există diferențe semnificative de vârstă între băieți și fete. Sursa: elaborare proprie.

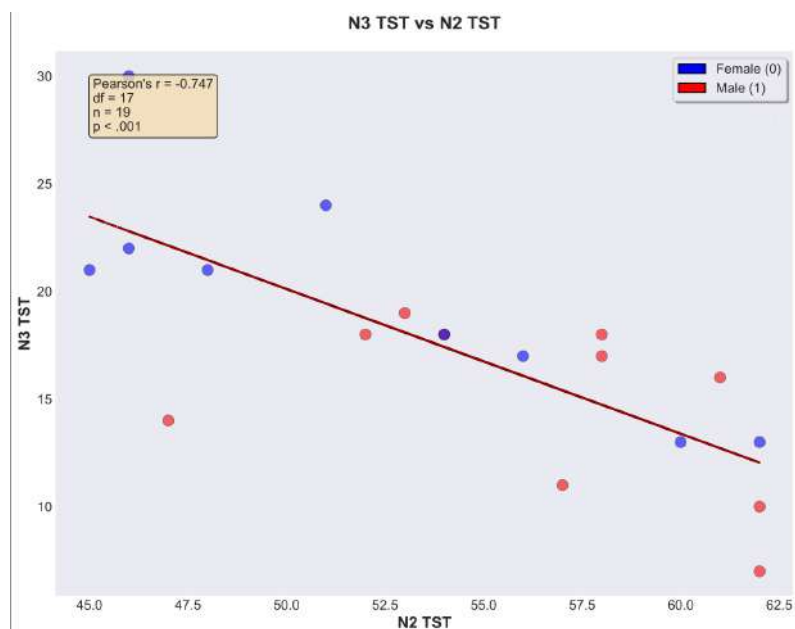


Figura 4.3 Relația dintre proporțiile stadiilor N3 și N2 din TST. Corelația negativă observată reflectă o tendință de compensare reciprocă a celor două stadii NREM, cu implicații pentru distribuția de clase în antrenare. Sursa: elaborare proprie.

Capitolul 5

Metodologie

Pipeline-ul experimental urmează patru etape secvențiale: pregătirea datelor, pre-antrenarea pe date adulte, *adaptarea* pe date pediatrice și evaluarea comparativă.

5.1 Preprocesarea semnalelor

Preprocesarea este orchestrată de un script centralizat care apelează module specializate pentru fiecare combinație (dataset, model). Pașii comuni sunt:

1. **Citirea semnalului brut:** fișiere EDF (HMC, Sleep-EDF) sau format NPA pacient (setul pediatric), cu extragerea canalului EEG relevant.
2. **Conversia adnotărilor:** stadiile din formatul original (R&K pentru Sleep-EDF [22], AASM text pentru HMC, JSON pentru setul pediatric) sunt mapate în schema canonică cu 5 clase: W=0, N1=1, N2=2, N3=3, REM=4.
3. **Segmentarea în epoci de 30 s:** semnalul continuu este decupat în ferestre non-suprapuse de 30 de secunde aliniate la momentele de start ale adnotărilor. Epocile în afara limitelor semnalului sunt eliminate.
4. **Reeșantionarea:** fiecare epocă este reeșantionată la rata țintă a modelului: **100 Hz** pentru AttnSleep și DeepSleepNet (3000 eșantioane/epocă), respectiv **256 Hz** pentru PediaSleepNet (7680 eșantioane/epocă).
5. **Salvarea:** fiecare înregistrare produce un fișier binar comprimat conținând matricea de epoci $\mathbf{X} \in \mathbb{R}^{N \times T}$, vectorul de etichete $\mathbf{y} \in \{0, \dots, 4\}^N$, rata de eșantionare și metadata de trasabilitate.

5.1.1 Preprocesarea fiecărui dataset

HMC

Semnalul este citit din canalul EEG F4-M1. Această derivație a fost aleasă deoarece surprinde activitatea central-frontală relevantă atât pentru fuzurile de somn (maxime

frontal) cât și pentru undele delta, asigurând compatibilitate cu cerințele standardului AASM [12]. Se rețin doar epocile cu durată de exact 30 s; epocile parțiale de la limitele înregistrării sunt eliminate. Din cele 151 de înregistrări disponibile, **148 de pacienți** au intrat în experimentele de pre-antrenare (3 înregistrări excluse din cauza erorilor de semnal).

Sleep-EDF

Fișierele PSG sunt asociate cu hipnogramele corespunzătoare prin interfața MNE. S-a utilizat canalul EEG **Fpz-Cz**, consistent cu canalul folosit în lucrările de referință [6, 25], asigurând reproducibilitatea comparațiilor. Adnotările cu durată multiplă de 30 s sunt expandate în epoci consecutive, deoarece unele etichete din formatul R&K [22] acoperă mai multe epoci consecutive; expansiunea produce structura de o-etichetă-per-epocă necesară antrenării.

Setul de date pediatric

Fișierele EEG sunt stocate în format NPA per pacient. Pacienții pentru care decodarea eșuează sunt raportați într-un jurnal de erori și excluși din pipeline, astfel că eșecurile izolate nu afectează procesarea celorlalte înregistrări. Semnalul EEG este extras din derivația bipolară **F4-C4** (diferența semnalelor de la electrozii F4 și C4), utilizată uniform pentru AttnSleep, DeepSleepNet și PediaSleepNet. Această derivație central-frontală surprinde activitatea pentru stadiile de somn și menține un montaj central consistent cu selecțiile din dataseturile adulte. Stadiile din fișierele de hipnogramă JSON sunt normalizate la schema canonică cu 5 clase.

Pentru PediaSleepNet, pe lângă semnalul EEG și etichete, se propagă și metadatele de vârstă: vârsta pacientului și grupa demografică ($G0$ sau $G1$), necesare filtrelor adaptive depthwise și modulelor de condiționare FiLM.

5.2 Antrenarea modelelor (pre-antrenare pe date adulte)

Pre-antrenarea s-a realizat separat pe fiecare combinație model $\times$ dataset adult (HMC și Sleep-EDF), rezultând opt experimente independente.

5.2.1 DeepSleepNet

Antrenarea DeepSleepNet urmează strategia în două faze descrisă în [25]:

Faza 1: Pre-antrenare CNN (DeepFeatureNet). Cele două rețele CNN sunt antrenate pe epoci individuale, cu supraeșantionarea claselor minoritare prin *random oversampling* pentru echilibrarea distribuțiilor. Se utilizează optimizatorul Adam [16] cu $lr = 1 \times 10^{-4}$, $\beta_1 = 0,9$, $\beta_2 = 0,999$ și *dropout* 0,5 [24]. Dimensiunea batch-ului este de 100 de epoci. Intrarea este semnalul EEG monocanal reeșantionat la 100 Hz (3000 de eșantioane per epocă). Modelul are aproximativ 22,6 milioane de parametri antrenabili.

Faza 2: Antrenare secvențială. Modelul complet (CNN + BiLSTM) este antrenat pe secvențe de 25 de epoci consecutive, cu *batch* de 10 secvențe. Stratul CNN utilizează o rată de învățare redusă (1×10^{-6}) pentru a proteja trăsăturile pre-antrenate, în timp ce straturile BiLSTM utilizează 1×10^{-4} , cu *gradient clipping* la valoarea 10. Antrenarea se efectuează pe 200 de epoci, cel mai bun model fiind selectat după *macro F1-score*-ul de validare.

5.2.2 AttnSleep

Antrenarea AttnSleep urmează protocolul original [6]. Intrarea este semnalul EEG monocanal la 100 Hz (3000 de eșantioane per epocă).

Nucleul mic al MRCNN are dimensiunea 50 de eșantioane (fereastră de 0,5 s, vizând frecvențele alfa/theta de 4–13 Hz), iar nucleul mare are dimensiunea 400 de eșantioane (fereastră de 4 s, vizând delta <4 Hz). Modulul AFR utilizează *Squeeze-and-Excitation* cu $r = 16$ [11]. TCE conține 2 blocuri stivuite cu 8 capete de atenție, $d_{\text{model}} = 128$ și convoluție depthwise cauzală cu nucleu 7.

Funcția de pierdere este cross-entropy cu ponderi logaritmice de clasă [6]. Optimizatorul este Adam ($\beta_1 = 0,9$, $\beta_2 = 0,999$, *weight decay* = 1×10^{-3}) cu $lr = 1 \times 10^{-3}$ pentru primele 10 epoci, redus la 1×10^{-4} ulterior. Dimensiunea batch-ului: 32. *Early stopping* cu răbdare de 10 epoci pe pierderea de validare; numărul maxim de epoci: 100. Modelul are aproximativ 523.000 de parametri antrenabili.

Protocolul de validare este partiția stratificată 70/10/20 (antrenare/validare/test).

5.2.3 PediaSleepNet

Pre-antrenarea se realizează separat pentru **PediaSleepNet Base** și **PediaSleepNet V2** pe fiecare dataset adult, rezultând patru experimente independente ($\text{Base} \times \{\text{HMC}, \text{Sleep-EDF}\}$ și $\text{V2} \times \{\text{HMC}, \text{Sleep-EDF}\}$).

Intrare: epoci EEG la 256 Hz (7680 de eșantioane per epocă), cu metadata de vârstă implicite pentru adulți (grupă = $G1$) și embedding de vârstă de 64 de dimensiuni, inițializat uniform.

MRCNN+ folosește trei ramuri CNN paralele: nucleu mic $k = 50$ cu pas 6, nucleu mediu $k = 150$ cu pas 12 și nucleu mare $k = 400$ cu pas 50. Blocul SE adaptiv pe vârstă utilizează $r = 16$ [11]. Straturile FiLM sunt aplicate după MRCNN+ și după fiecare bloc TCE, cu γ și β inițializate la 0. TCE are 2 blocuri, 4 capete de atenție, $d_{\text{model}} = 128$, dimensiunea ascunsă FFN = 192 și $\text{dropout} = 0,5$.

Diferența dintre cele două versiuni la pre-antrenare constă exclusiv în modulul MRCNN+: Base concatenează temporal cele trei ramuri înainte de proiecția 1×1 comună, în timp ce V2 proiectează individual fiecare ramură la $d_{\text{model}} = 128$ și le combină prin medie element-wise. V2 include totodată capul auxiliar N1 antrenat prin pierdere BCE secundară.

Funcția de pierdere principală este cross-entropy cu *label smoothing* ($\varepsilon = 0,1$), cost de tranziție (pondere 0,06) și regularizare temporală KL (pondere 0,02). Optimizatorul este Adam cu $\text{lr} = 1 \times 10^{-3}$ și $\text{weight decay} = 1 \times 10^{-4}$, batch de 128 de secvențe, maximum 100 de epoci. PediaSleepNet Base are aproximativ 1,1 milioane de parametri antrenabili, iar V2 aproximativ 945.000.

Protocolul de validare este partiția stratificată 70/10/20 (antrenare/validare/test).

5.3 Adaptarea pe date pediatrice

Adaptarea urmează un protocol de validare încrucișată pe 5 folduri pe setul pediatric. La fiecare fold $k \in \{1, \dots, 5\}$, aproximativ 16 pacienți sunt utilizați ca set de antrenare iar 4 ca set de validare. Modelul este inițializat cu *checkpoint*-ul pre-antrenat pe adulți, iar criteriul de selecție al celui mai bun model este *macro F1-score* pe validare.

Adaptarea a fost efectuat pentru **toate cele patru modele** pe ambele seturi de pre-antrenare, rezultând $8 \text{ experimente} \times 5 \text{ folduri} = 40$ de antrenări independente:

- AttnSleep (HMC - pediatric)
- AttnSleep (Sleep-EDF - pediatric)
- DeepSleepNet (HMC - pediatric)
- DeepSleepNet (Sleep-EDF - pediatric)
- PediaSleepNet Base (HMC - pediatric)
- PediaSleepNet Base (Sleep-EDF - pediatric)
- PediaSleepNet V2 (HMC - pediatric)

- **PediaSleepNet V2 (Sleep-EDF - pediatric)**

Fiecare experiment produce: un *checkpoint* al celui mai bun model; metricile de evaluare per fold (acuratețe, *macro F1-score*, Cohen's κ , *F1-score* per clasă); și o vizualizare a matricei de confuzie. Metricile agregate (medie $\pm$ abatere standard pe cele 5 folduri) sunt salvate la finalul fiecărui experiment.

5.3.1 Strategia de *adaptare*

Strategia urmează principiile clasice ale *transfer learning*, adaptate pentru clasificarea somnului. Indiferent de model, două principii rămân comune. **(i) Protejarea straturilor de extracție de trăsături.** Reprezentările de nivel scăzut ale semnalului EEG (detectori de unde delta, fusuri de somn, complexe-K) învățate pe seturile adulte rămân relevante și pentru copii; ele sunt protejate în timpul adaptării, fie prin înghețare, fie printr-o rată de învățare mult mai redusă, pentru a preveni *catastrophic forgetting*, riscul principal atunci când setul de *adaptare* este mic (19 pacienți). **(ii) Selecția pe *macro F1-score* de validare.** Cel mai bun *checkpoint* al fiecărui fold este ales după *macro F1-score*-ul de validare, nu după pierdere, dat fiind dezechilibrul accentuat al claselor.

Hiperparametrii concreți (ratele de învățare, numărul de epoci, straturile actualizate) diferă însă de la model la model, urmând specificul fiecărei arhitecturi.

DeepSleepNet

Adaptarea modelului DeepSleepNet reia strategia în două faze din pre-antrenare. Modelul complet (CNN + BiLSTM) este inițializat din *checkpoint*-ul DeepFeatureNet pre-antrenat pe adulți (se reîncarcă ponderile CNN, excluzând parametrii optimizatorului și ai capului softmax), după care întreaga rețea este antrenată pe secvențe de 25 de epoci consecutive, cu *batch* de 10 secvențe. Nu se aplică o înghețare explicită a straturilor; protejarea trăsăturilor CNN se realizează prin **rate de învățare diferențiate**: straturile convoluționale pre-antrenate utilizează $lr = 1 \times 10^{-6}$, iar straturile BiLSTM și capul de clasificare nou inițializate utilizează $lr = 1 \times 10^{-4}$, cu *gradient clipping* la valoarea 10 pentru stabilitate. Numărul de epoci este de 200.

AttnSleep

AttnSleep transferă blocurile MRCNN, AFR și TCE pre-antrenate, reantrenându-le pe datele pediatrie cu o rată de învățare redusă față de pre-antrenare și păstrând funcția de pierdere cross-entropy cu ponderi logaritmice de clasă [6]. Selecția modelului și oprirea timpurie urmează *macro F1-score*-ul de validare pe fiecare fold.

PediaSleepNet Base și V2

Pentru PediaSleepNet Base și V2, *adaptarea* urmează un protocol de dezghețare progresivă în trei etape: **E1** ($lr = 5 \times 10^{-4}$) dezghează capetele de clasificare, encodarea pozițională și, pentru V2, capul auxiliar N1; **E2** ($lr = 2 \times 10^{-4}$) adaugă straturile de proiecție, SE și FiLM; **E3** ($lr = 5 \times 10^{-5}$) dezghează modelul complet. Filtrele adaptive depthwise și modulele FiLM, care la pre-antrenare operau cu grupa implicită $G1$ (adulți), sunt acum actualizate pe ambele grupe demografice ($G0$ și $G1$) prezente în populația pediatrică (vârste 9–17 ani). Augmentarea datelor (zgomot gaussian, scalare amplitudine, deplasare temporală) este activă pe durata întregii etape de adaptare.

Validare încrucișată pe 5 folduri. Pentru toate modelele, cei 19 de pacienți pediatrici sunt împărțiți în 5 folduri stratificate (aproximativ 16 pacienți pentru antrenare și 4 pentru validare la fiecare fold), asigurând că fiecare pacient apare cel puțin o dată în setul de validare.

5.4 Mediul software și hardware

Toate experimentele au fost executate pe **Google Colab** cu un GPU **NVIDIA H100**. Colab oferă mașini virtuale Linux cu drivere CUDA preinstalate și acces la Google Drive pentru persistarea *checkpoint*-urilor între sesiuni.

Tabelul 5.1 prezintă configurația hardware, iar Tabelul 5.2 detaliază bibliotecile software utilizate.

Componentă	Detalii
GPU	NVIDIA H100 (80 GB HBM3)
Platformă	Google Colab Pro
Stocare	Google Drive (persistare <i>checkpoint</i> -uri)

Tabel 5.1 Configurația hardware și platforma de execuție.

Bibliotecă	Versiune	Scop
Python	3.10	Limbaaj principal de programare
PyTorch [20]	≥ 2.0	Antrenarea AttnSleep și PediaSleepNet
TensorFlow	2.x (com- pat.v1)	Antrenarea DeepSleepNet
NumPy	≥ 1.23	Operații numerice pe array-uri
SciPy	≥ 1.9	Filtrare semnal, statistici, reeșantionare
scikit-learn	≥ 1.1	Metrici de evaluare ($F1$ -score, κ , confusion matrix), stratificare folduri
MNE	≥ 1.0	Citire fișiere EDF, preprocesare EEG
pyEDFlib		Citire alternativă fișiere EDF/EDF+
matplotlib	≥ 3.5	Vizualizare confusion matrix
seaborn		Vizualizare statistică (heatmaps, distribuții)
pandas	≥ 1.4	Metadate pacienți, tabele de rezultate
joblib		Paralelizare preprocesare
CUDA / cuDNN	12.x	Accelerare GPU
Google Colab		Platformă cloud de execuție
Google Drive API		Persistare <i>checkpoint</i> -uri

Tabel 5.2 Mediul software complet utilizat în proiect. Versiunile indicate sunt cele minime testate.

5.5 Metrici de evaluare

Performanța modelelor este evaluată printr-un set de metrici, alese pentru a oferi atât o imagine globală cât și o analiză detaliată per clasă, esențială în contextul dezechilibrului pronunțat al distribuției stadiilor de somn.

5.5.1 Acuratețe globală (*Accuracy*)

Acuratețea globală reprezintă proporția de epoci clasificate corect din totalul epocilor evaluate:

$$\text{ACC} = \frac{\sum_{k=1}^K \text{TP}_k}{N}, \quad (5.1)$$

unde TP_k (*true positives*) este numărul de epoci din clasa k clasificate corect, $K = 5$ este numărul de clase, iar N este numărul total de epoci evaluate. Deși intuitivă, acuratețea este o metrică înșelătoare în contextul dezechilibrului de clase din somnul pediatric: stadiul N2 reprezintă $\sim 54\%$ din TST, astfel că un model care clasifică toate epocile ca N2 ar obține acuratețe de 54% fără utilitate clinică. Acuratețea este folosită ca indicator complementar,

nu ca metrică primară de comparație.

5.5.2 *F1-score* per clasă

F1-score-ul reprezintă media armonică dintre precizie (*Precision*) și rata de regăsire (*Recall*), fiind calculat independent pentru fiecare clasă k :

$$F1_k = \frac{2 \cdot \text{Precision}_k \cdot \text{Recall}_k}{\text{Precision}_k + \text{Recall}_k}, \quad \text{Precision}_k = \frac{\text{TP}_k}{\text{TP}_k + \text{FP}_k}, \quad \text{Recall}_k = \frac{\text{TP}_k}{\text{TP}_k + \text{FN}_k}, \quad (5.2)$$

unde TP_k (*true positives*) este numărul de epoci din clasa k clasificate corect, FP_k (*false positives*) este numărul de epoci atribuite eronat clasei k , iar FN_k (*false negatives*) este numărul de epoci din clasa k clasificate eronat în alte clase. Precizia (Precision_k) măsoară proporția predicțiilor corecte din totalul epocilor atribuite clasei k , rata de regăsire (Recall_k) măsoară proporția epocilor reale din clasa k identificate corect, iar $F1_k \in [0, 1]$ echilibrează cele două prin media lor armonică, penalizând modelele care optimizează doar una dintre ele.

5.5.3 *Macro F1-score*

Macro F1-score este media aritmetică a *F1-score*-urilor calculate independent pentru fiecare clasă:

$$\text{Macro-F1} = \frac{1}{K} \sum_{k=1}^K F1_k, \quad (5.3)$$

unde $K = 5$ este numărul de clase, iar $F1_k$ este *F1-score*-ul clasei k definit anterior. Această metrică ponderează egal clasele indiferent de frecvența lor, astfel că performanța pe stadiul N1 ($\sim 9\%$ din TST) are aceeași semnificație ca performanța pe N2 ($\sim 54\%$). Din aceste motive, *macro F1-score* este metrica principală de comparație, în concordanță cu practica din literatura de referință [6, 25].

5.5.4 Cohen's κ

Coeфициntul Cohen's κ [5] măsoară corespondența dintre predicțiile modelului și etichetele expertului, corectată pentru corespondența aleatoare:

$$\kappa = \frac{p_o - p_e}{1 - p_e}, \quad p_e = \sum_{k=1}^K \frac{n_{k,\text{true}} \cdot n_{k,\text{pred}}}{N^2} \quad (5.4)$$

unde p_o este acuratețea observată (proporția epocilor pe care modelul și expertul le clasifică identic), p_e este concordanța așteptată întâmplător, $n_{k,\text{true}}$ și $n_{k,\text{pred}}$ sunt numărul de epoci

din clasa k conform etichetelor expertului respectiv predicțiilor modelului, $K = 5$ este numărul de clase, iar N este numărul total de epoci. κ este uzual utilizat în clinică pentru evaluarea fiabilității inter-evaluator, iar aplicarea sa la comparația model vs. expert plasează performanța modelului pe o scală cu semnificație clinică directă [12]. Concordanța inter-expert pentru clasificarea somnului se situează în intervalul 0,76–0,82, astfel că un model cu $\kappa > 0,75$ poate fi considerat comparabil cu un expert uman.

5.5.5 Matricea de confuzie

Matricea de confuzie 5×5 completează această analiză relevând tipare de eroare sistematice: N1 confundat cu W sau N2; N3 confundat cu N2 la granițele de stadiu; REM confundat cu N1 sau W din cauza activității EEG de amplitudine mică. Aceste tipare permit interpretarea directă a limitelor transferului între domenii adult-pediatric.

Capitolul 6

Rezultate și evaluare

Acest capitol prezintă, pentru fiecare dintre cele patru modele, rezultatele obținute în cele două etape ale pipeline-ului: **pre-antrenarea** pe seturile adulte (Sleep-EDF și HMC) și **adaptarea** pe setul de date pediatric. Rezultă $4 \times 2 = 8$ experimente de pre-antrenare și 8 experimente de *adaptare*, raportate individual în secțiunile 6.1 și 6.2. Pentru fiecare experiment se prezintă metricile globale (acuratețe ACC, *macro F1-score* MF1 și Cohen's κ), *F1-score*-ul per clasă și matricea de confuzie normalizată pe rânduri.

Trebuie subliniat că protocoalele de evaluare diferă între cele două etape: pre-antrenarea este evaluată pe partițiile interne ale fiecărui set adult, în timp ce *adaptarea* este evaluat prin validare încrucișată pe 5 folduri pe setul pediatric (medie $\pm$ abatere standard). Din acest motiv, valorile din acest capitol descriu comportamentul fiecărui model pe domeniul său de evaluare; **compararea directă între modele și sinteza concluziilor sunt realizate în Capitolul 7**. Metrica primară de referință rămâne *macro F1-score*-ul.

6.1 Rezultate la pre-antrenare

Pre-antrenarea a urmat, pentru fiecare model, protocolul de validare propriu: DeepSleepNet a fost evaluat pe 3 folduri de tip *leave-subjects-out*; AttnSleep pe o partiție stratificată 70/10/20 pe Sleep-EDF, respectiv *leave-one-subject-out* pe cei 22 de subiecți de test ai HMC; iar PediaSleepNet Base și V2 pe o partiție stratificată 70/10/20. Valorile pentru PediaSleepNet și pentru AttnSleep–HMC reprezintă rezultate agregate pe întregul set de test, fără abatere standard între folduri.

6.1.1 DeepSleepNet (Sleep-EDF)

Pe Sleep-EDF, DeepSleepNet atinge $ACC = 89,3\%$ ($\pm 3,4$), $MF1 = 74,0\%$ ($\pm 4,9$) și $\kappa = 0,80$ (Tabelul 6.1). Stadiile W, N2 și N3 sunt clasificate robust, însă N1 rămâne dificil ($F1 = 36,4\%$), fiind confundat preponderent cu N2 și REM, tipar vizibil în matricea de confuzie din Figura 6.1.

ACC	MF1	κ	W	N1	N2	N3	REM
89,3	74,0	0,80	96,4	36,4	86,8	83,6	67,0

Tabel 6.1 DeepSleepNet pre-antrenat pe Sleep-EDF: metrice globale și $F1$ -score per clasă (%).

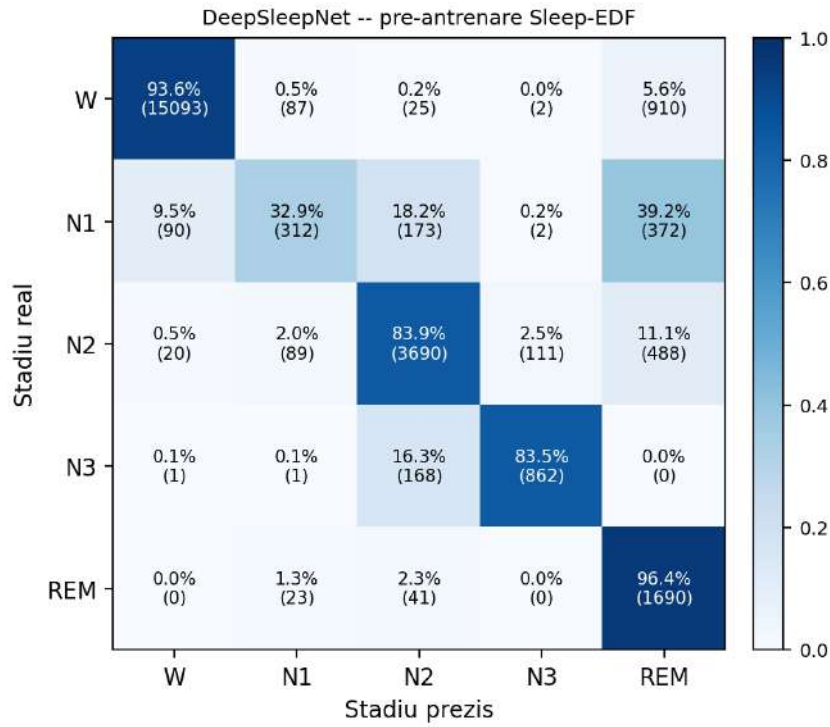


Figura 6.1 Matricea de confuzie normalizată – DeepSleepNet pre-antrenat pe Sleep-EDF. Sursa: elaborare proprie.

6.1.2 DeepSleepNet (HMC)

Pe HMC, performanța DeepSleepNet scade semnificativ și este foarte instabilă între folduri: $ACC = 55,5\%$ ($\pm 13,0$), $MF1 = 54,6\%$ ($\pm 13,7$) și $\kappa = 0,44$ (Tabelul 6.2). Confuzia masivă $N2 \leftrightarrow N3$ ($F1$ pentru $N2$ doar $42,1\%$) și variabilitatea ridicată indică o adaptare slabă la populația clinică heterogenă a HMC (Figura 6.2).

ACC	MF1	κ	W	N1	N2	N3	REM
55,5	54,6	0,44	74,5	31,2	42,1	56,6	68,9

Tabel 6.2 DeepSleepNet pre-antrenat pe HMC: metrice globale și $F1$ -score per clasă (%).

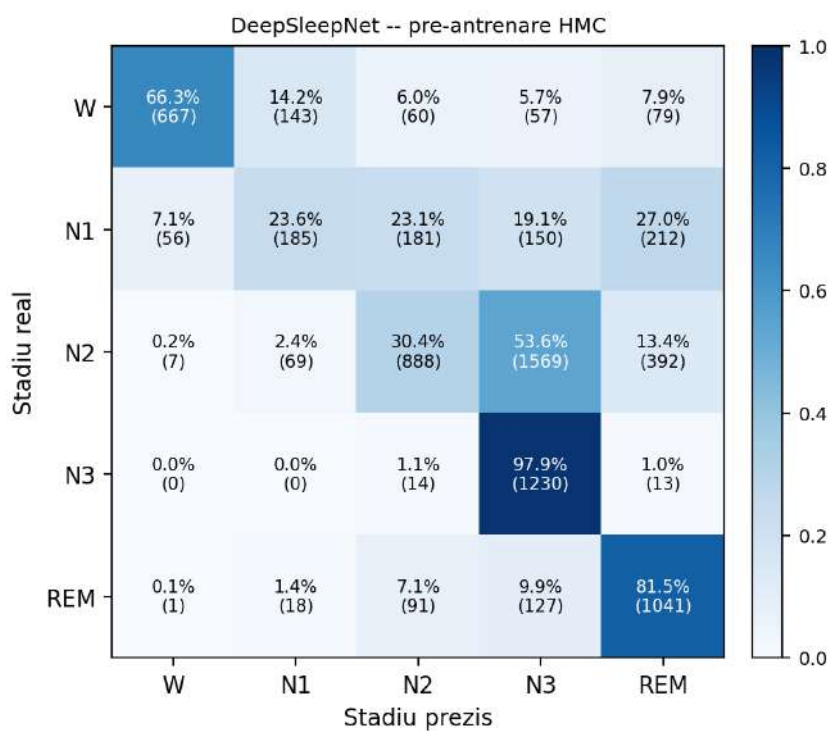


Figura 6.2 Matricea de confuzie normalizată – DeepSleepNet pre-antrenat pe HMC. Sursa: elaborare proprie.

6.1.3 AttnSleep (Sleep-EDF)

AttnSleep obține cele mai bune rezultate la pre-antrenare pe Sleep-EDF: $ACC = 95,5\%$, $MF1 = 80,5\%$ și $\kappa = 0,89$ (Tabelul 6.3). Toate stadiile, cu excepția N1 (42,7%), depășesc 78% $F1-score$, iar matricea de confuzie (Figura 6.3) este puternic diagonală.

ACC	MF1	κ	W	N1	N2	N3	REM
95,5	80,5	0,89	99,2	42,7	89,1	92,4	78,9

Tabel 6.3 AttnSleep pre-antrenat pe Sleep-EDF: metrice globale și $F1-score$ per clasă (%).

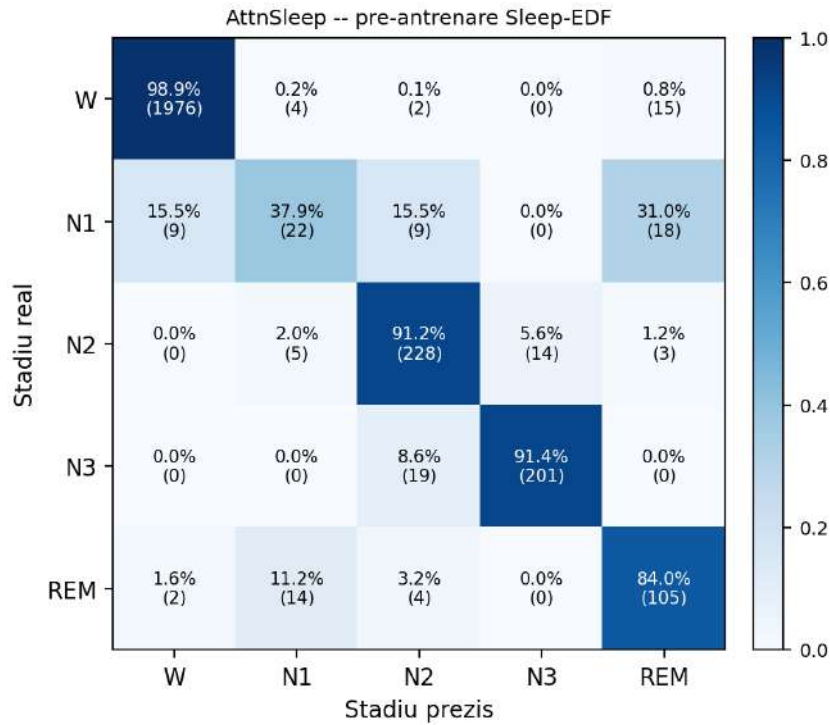


Figura 6.3 Matricea de confuzie normalizată – AttnSleep pre-antrenat pe Sleep-EDF. Sursa: elaborare proprie.

6.1.4 AttnSleep (HMC)

Pe HMC, AttnSleep rămâne competitiv (agregat pe 22 de subiecți de test): $ACC = 85,3\%$, $MF1 = 78,1\%$ și $\kappa = 0,80$ (Tabelul 6.4), cu cel mai bun $F1$ -score pentru N1 dintre toate experimentele de pre-antrenare (44,6%). Matricea de confuzie este prezentată în Figura 6.4.

ACC	MF1	κ	W	N1	N2	N3	REM
85,3	78,1	0,80	95,5	44,6	81,5	86,5	82,4

Tabel 6.4 AttnSleep pre-antrenat pe HMC: metrice globale și $F1$ -score per clasă (%).

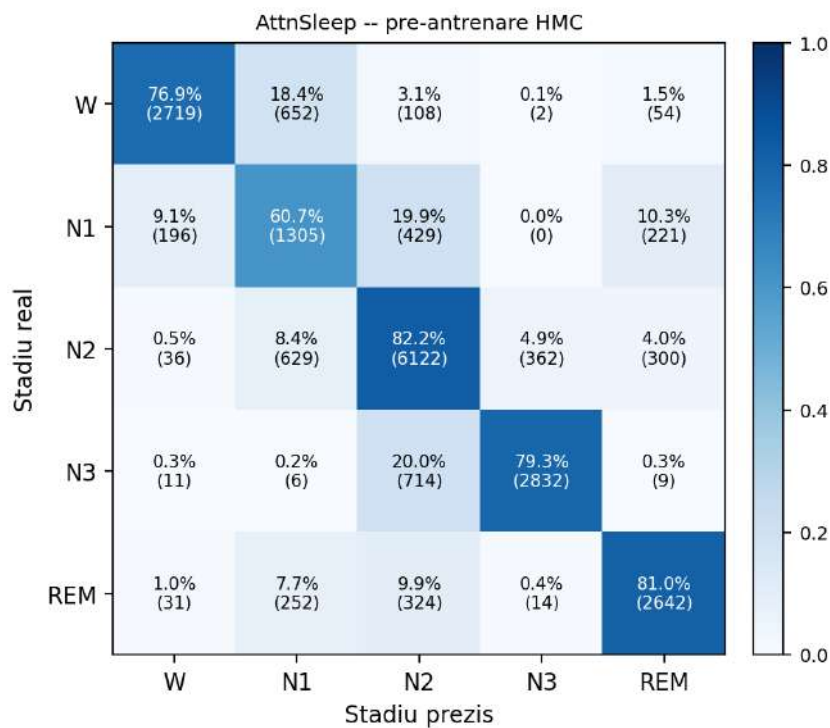


Figura 6.4 Matricea de confuzie normalizată – AttnSleep pre-antrenat pe HMC. Sursa: elaborare proprie.

6.1.5 PediaSleepNet Base (Sleep-EDF)

PediaSleepNet Base obține pe Sleep-EDF $ACC = 87,2\%$, $MF1 = 75,0\%$ și $\kappa = 0,77$ (Tabelul 6.5), cu un $F1-score$ pentru N1 (42,2%) comparabil cu al AttnSleep, în ciuda arhitecturii compacte ($\sim 1,1$ M parametri). Matricea de confuzie este prezentată în Figura 6.5.

ACC	MF1	κ	W	N1	N2	N3	REM
87,2	75,0	0,77	96,6	42,2	79,9	83,2	73,1

Tabel 6.5 PediaSleepNet Base pre-antrenat pe Sleep-EDF: metrice globale și $F1-score$ per clasă (%).

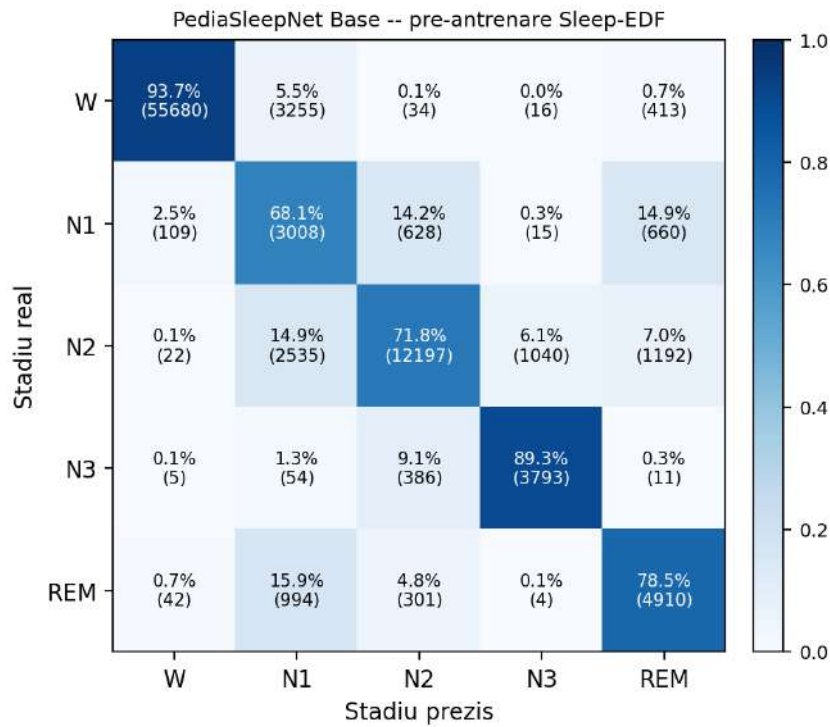


Figura 6.5 Matricea de confuzie normalizată – PediaSleepNet Base pre-antrenat pe Sleep-EDF. Sursa: elaborare proprie.

6.1.6 PediaSleepNet Base (HMC)

Pe HMC, PediaSleepNet Base atinge $ACC = 71,7\%$, $MF1 = 70,3\%$ și $\kappa = 0,64$ (Tabelul 6.6). Deși metricile globale sunt sub cele ale AttnSleep, modelul înregistrează cel mai ridicat $F1-score$ pentru N1 din întreaga etapă de pre-antrenare (46,1%) (Figura 6.6).

ACC	MF1	κ	W	N1	N2	N3	REM
71,7	70,3	0,64	78,4	46,1	71,9	81,4	73,6

Tabel 6.6 PediaSleepNet Base pre-antrenat pe HMC: metrici globale și $F1-score$ per clasă (%).

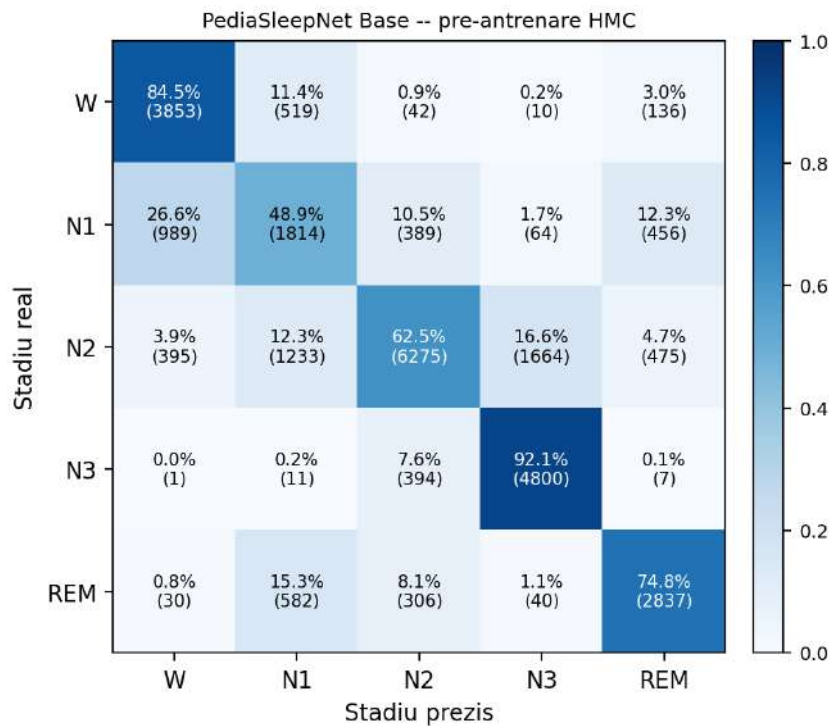


Figura 6.6 Matricea de confuzie normalizată – PediaSleepNet Base pre-antrenat pe HMC. Sursa: elaborare proprie.

6.1.7 PediaSleepNet V2 (Sleep-EDF)

PediaSleepNet V2 obține pe Sleep-EDF rezultate aproape identice cu versiunea Base: ACC = 86,9%, MF1 = 74,5% și $\kappa = 0,77$ (Tabelul 6.7), cu cel mai bun *F1-score* pentru N1 dintre modelele PediaSleepNet pe acest set (43,4%), efect al capului auxiliar N1 (Figura 6.7).

ACC	MF1	κ	W	N1	N2	N3	REM
86,9	74,5	0,77	97,3	43,4	76,3	81,8	73,6

Tabel 6.7 PediaSleepNet V2 pre-antrenat pe Sleep-EDF: metrice globale și *F1-score* per clasă (%).

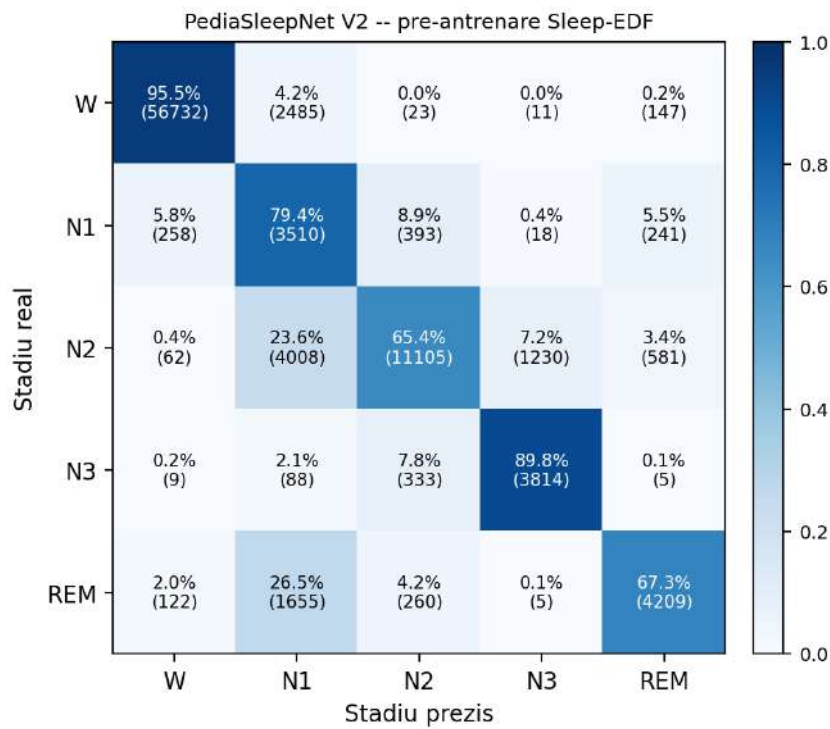


Figura 6.7 Matricea de confuzie normalizată – PediaSleepNet V2 pre-antrenat pe Sleep-EDF. Sursa: elaborare proprie.

6.1.8 PediaSleepNet V2 (HMC)

Pe HMC, PediaSleepNet V2 atinge $ACC = 72,3\%$, $MF1 = 69,5\%$ și $\kappa = 0,64$ (Tabelul 6.8), valori comparabile cu cele ale versiunii Base (Figura 6.8).

ACC	MF1	κ	W	N1	N2	N3	REM
72,3	69,5	0,64	78,7	40,9	74,0	78,5	75,3

Tabel 6.8 PediaSleepNet V2 pre-antrenat pe HMC: metrici globale și $F1$ -score per clasă (%).

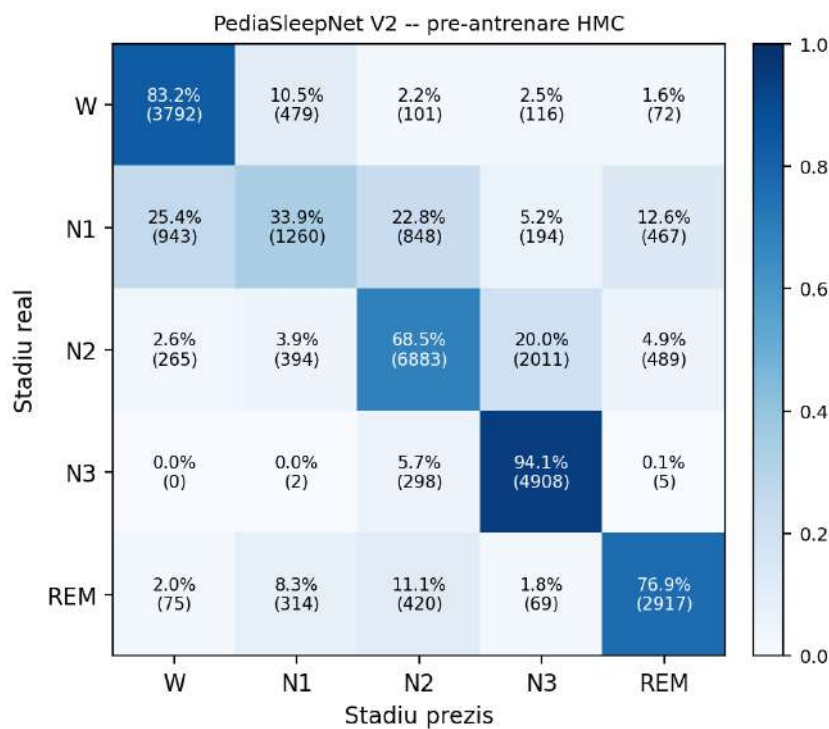


Figura 6.8 Matricea de confuzie normalizată – PediaSleepNet V2 pre-antrenat pe HMC. Sursa: elaborare proprie.

6.2 Rezultate la *adaptarea* pe date pediatrice

Această secțiune raportează rezultatele celor 8 experimente de *adaptare* pe setul pediatric, fiecare evaluat prin validare încrucișată pe 5 folduri. Valorile sunt prezentate ca medie pe cele 5 folduri, abaterea standard a metricilor globale fiind indicată în text. Matricele de confuzie reprezintă media normalizată pe cele 5 folduri.

6.2.1 DeepSleepNet (Sleep-EDF - pediatric)

Transferat de pe Sleep-EDF, DeepSleepNet obține cele mai slabe rezultate de pe setul pediatric: $ACC = 53,9\%$ ($\pm 8,8$), $MF1 = 37,1\%$ ($\pm 6,7$) și $\kappa = 0,32$ (Tabelul 6.9). Stadiul N1 este aproape complet ratat ($F1 = 7,5\%$), iar stadiul W este puternic confundat (Figura 6.9), semn al unei adaptări instabile pe setul pediatric mic.

ACC	MF1	κ	W	N1	N2	N3	REM
53,9	37,1	0,32	26,7	7,5	67,8	40,6	42,9

Tabel 6.9 DeepSleepNet (Sleep-EDF - pediatric): medie pe 5 folduri, $F1$ -score per clasă (%).

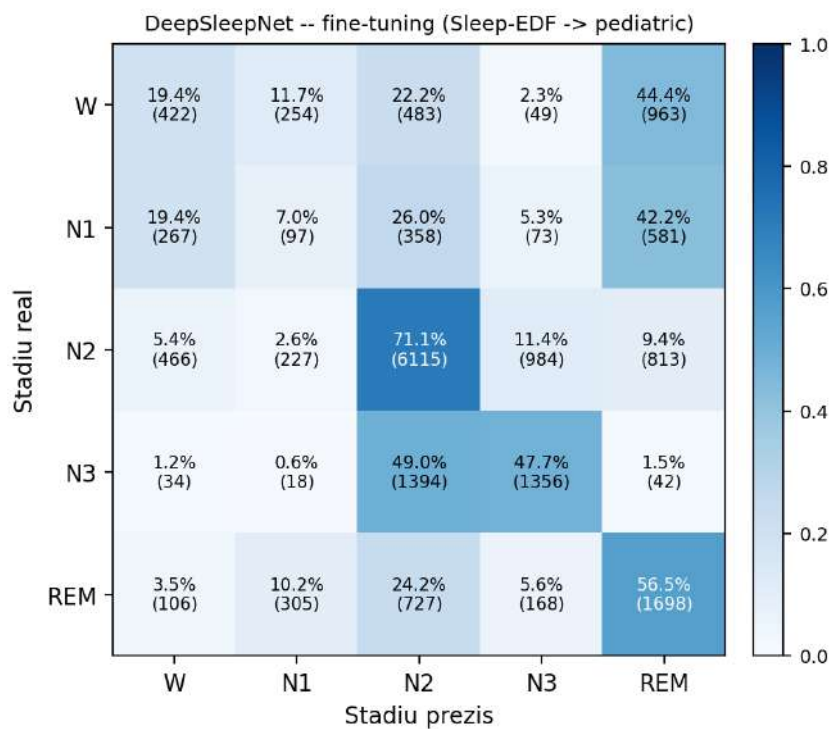


Figura 6.9 Matricea de confuzie normalizată – DeepSleepNet (Sleep-EDF - pediatric). Sursa: elaborare proprie.

6.2.2 DeepSleepNet (HMC - pediatric)

Transferat de pe HMC, DeepSleepNet se comportă ceva mai bine: $ACC = 58,4\%$ ($\pm 5,6$), $MF1 = 45,3\%$ ($\pm 7,4$) și $\kappa = 0,39$ (Tabelul 6.10). N1 rămâne clasa critică ($F1 = 12,3\%$), conform Figurii 6.10.

ACC	MF1	κ	W	N1	N2	N3	REM
58,4	45,3	0,39	48,1	12,3	66,5	52,3	47,5

Tabel 6.10 DeepSleepNet (HMC - pediatric): medie pe 5 folduri, $F1$ -score per clasă (%).

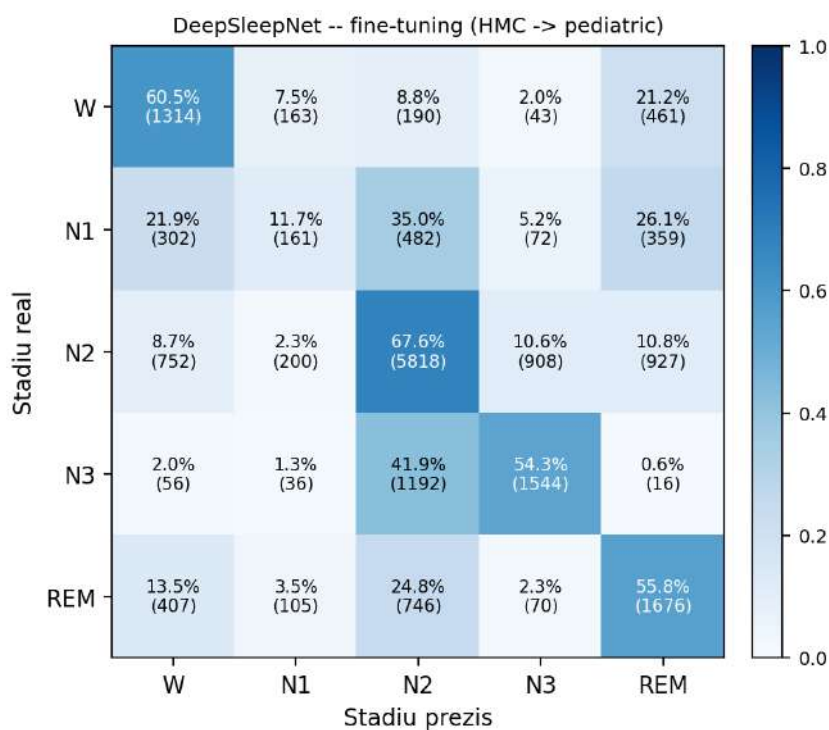


Figura 6.10 Matricea de confuzie normalizată – DeepSleepNet (HMC - pediatric). Sursa: elaborare proprie.

6.2.3 AttnSleep (Sleep-EDF - pediatric)

AttnSleep transferat de pe Sleep-EDF obține $ACC = 64,1\%$ ($\pm 5,0$), $MF1 = 53,4\%$ ($\pm 4,1$) și $\kappa = 0,48$ (Tabelul 6.11). În ciuda performanței ridicate la pre-antrenare, N1 scade în F1 pe datele pediatrice ($F1 = 13,4\%$), conform Figurii 6.11.

ACC	MF1	κ	W	N1	N2	N3	REM
64,1	53,4	0,48	52,4	13,4	73,5	66,1	61,8

Tabel 6.11 AttnSleep (Sleep-EDF - pediatric): medie pe 5 folduri, $F1$ -score per clasă (%).

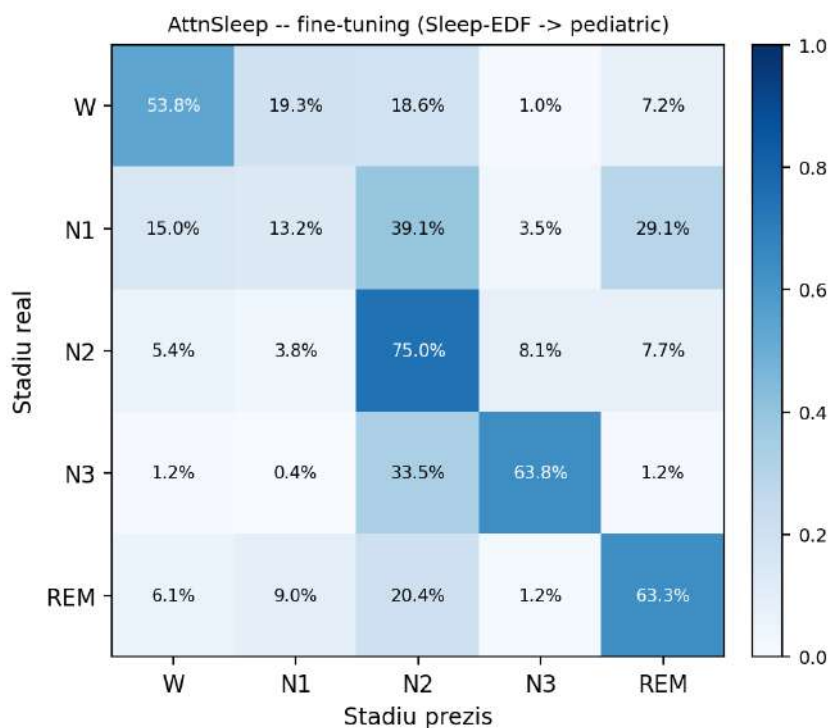


Figura 6.11 Matricea de confuzie normalizată – AttnSleep (Sleep-EDF - pediatric). Sursa: elaborare proprie.

6.2.4 AttnSleep (HMC - pediatric)

Cu backbone HMC, AttnSleep atinge $ACC = 67,7\%$ ($\pm 4,5$), $MF1 = 56,6\%$ ($\pm 3,6$) și $\kappa = 0,53$ (Tabelul 6.12), cel mai bun rezultat dintre modelele de referință pe setul pediatric, dar cu N1 tot scăzut (13,1%), conform Figurii 6.12.

ACC	MF1	κ	W	N1	N2	N3	REM
67,7	56,6	0,53	57,6	13,1	76,9	71,3	64,3

Tabel 6.12 AttnSleep (HMC - pediatric): medie pe 5 folduri, $F1$ -score per clasă (%).

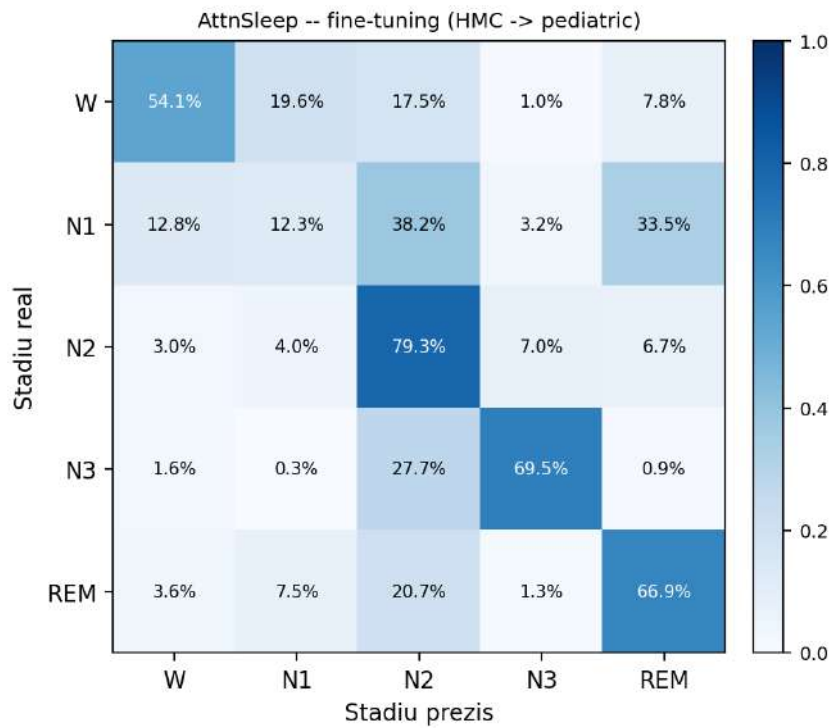


Figura 6.12 Matricea de confuzie normalizată – AttnSleep (HMC - pediatric). Sursa: elaborare proprie.

6.2.5 PediaSleepNet Base (Sleep-EDF - pediatric)

PediaSleepNet Base transferat de pe Sleep-EDF obține $ACC = 68,8\%$ ($\pm 5,6$), $MF1 = 61,6\%$ ($\pm 4,4$) și $\kappa = 0,57$ (Tabelul 6.13). $F1$ -score-ul pentru N1 (26,6%) este de aproape două ori mai mare decât al modelelor de referință pe acest backbone (Figura 6.13).

ACC	MF1	κ	W	N1	N2	N3	REM
68,8	61,6	0,57	63,7	26,6	74,2	72,4	71,1

Tabel 6.13 PediaSleepNet Base (Sleep-EDF - pediatric): medie pe 5 folduri, $F1$ -score per clasă (%).

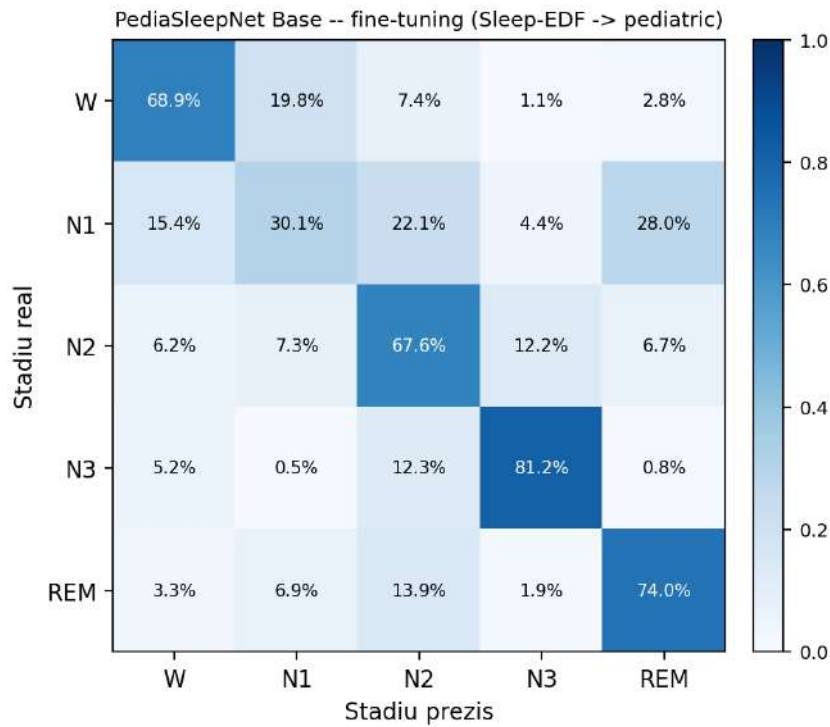


Figura 6.13 Matricea de confuzie normalizată – PediaSleepNet Base (Sleep-EDF - pediatric). Sursa: elaborare proprie.

6.2.6 PediaSleepNet Base (HMC - pediatric)

Cu backbone HMC, PediaSleepNet Base obține cel mai bun *macro F1-score* din întreaga lucrare pe setul pediatric: $ACC = 70,1\%$ ($\pm 4,9$), $MF1 = 64,6\%$ ($\pm 5,1$) și $\kappa = 0,59$ (Tabelul 6.14), cu N1 la 29,2% (Figura 6.14).

ACC	MF1	κ	W	N1	N2	N3	REM
70,1	64,6	0,59	78,3	29,2	73,7	71,9	69,9

Tabel 6.14 PediaSleepNet Base (HMC - pediatric): medie pe 5 folduri, *F1-score* per clasă (%).

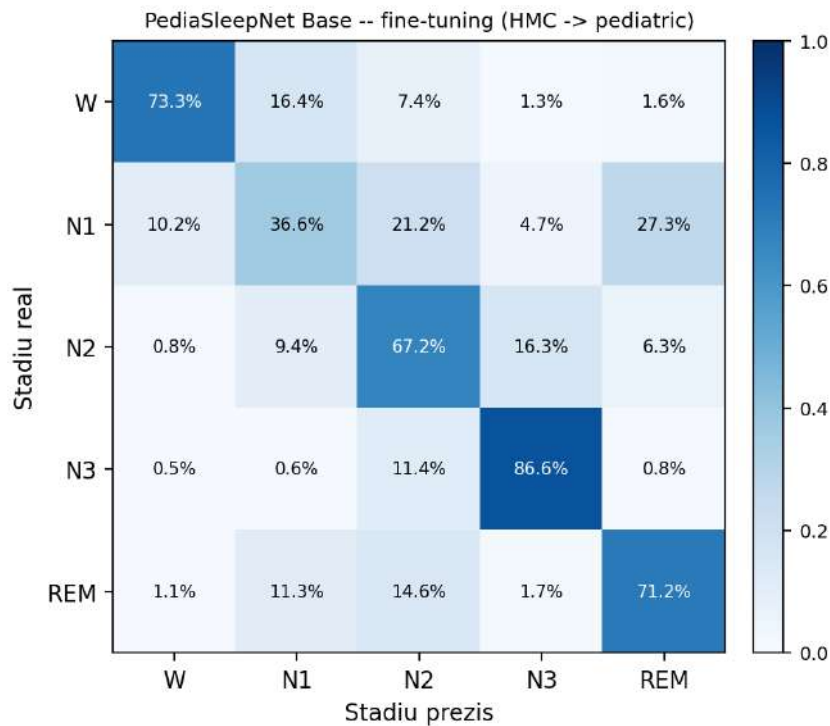


Figura 6.14 Matricea de confuzie normalizată – PediaSleepNet Base (HMC - pediatric). Sursa: elaborare proprie.

6.2.7 PediaSleepNet V2 (Sleep-EDF - pediatric)

PediaSleepNet V2 transferat de pe Sleep-EDF atinge cea mai bună acuratețe pe setul pediatric: $ACC = 71,2\%$ ($\pm 3,4$), $MF1 = 64,1\%$ ($\pm 3,9$) și $\kappa = 0,59$ (Tabelul 6.15). Capul auxiliar N1 ridică $F1$ -score-ul pentru N1 la 30,5%, cea mai bună valoare obținută pe acest backbone (Figura 6.15).

ACC	MF1	κ	W	N1	N2	N3	REM
71,2	64,1	0,59	72,7	30,5	75,8	71,7	69,9

Tabel 6.15 PediaSleepNet V2 (Sleep-EDF - pediatric): medie pe 5 folduri, $F1$ -score per clasă (%).

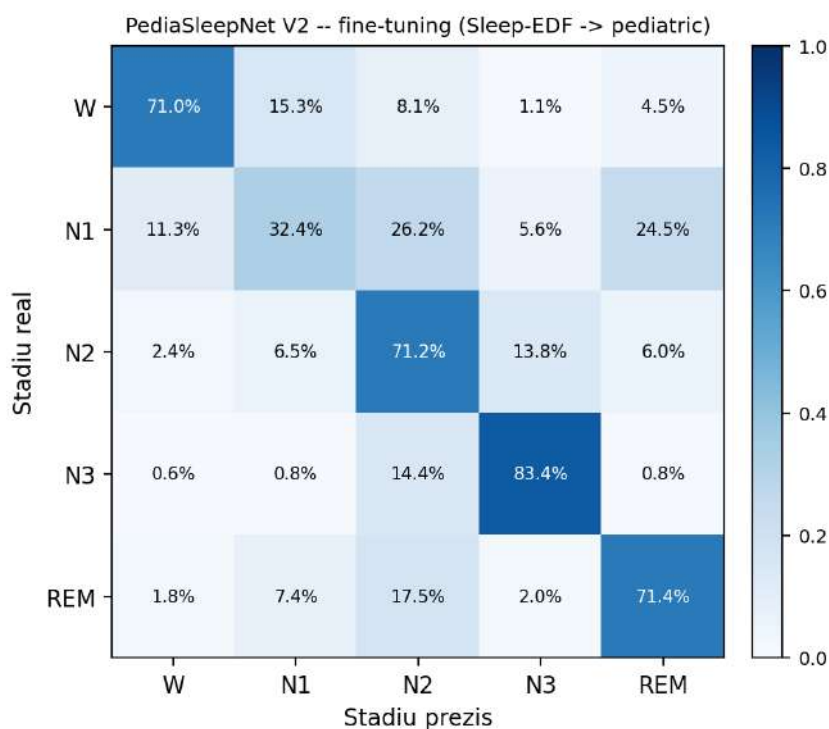


Figura 6.15 Matricea de confuzie normalizată – PediaSleepNet V2 (Sleep-EDF - pediatric). Sursa: elaborare proprie.

6.2.8 PediaSleepNet V2 (HMC - pediatric)

Cu backbone HMC, PediaSleepNet V2 obține $ACC = 67,7\%$ ($\pm 5,2$), $MF1 = 61,9\%$ ($\pm 5,1$) și $\kappa = 0,56$ (Tabelul 6.16), cu N1 la 28,2% (Figura 6.16).

ACC	MF1	κ	W	N1	N2	N3	REM
67,7	61,9	0,56	73,7	28,2	71,4	67,1	69,4

Tabel 6.16 PediaSleepNet V2 (HMC - pediatric): medie pe 5 folduri, *F1-score* per clasă (%).

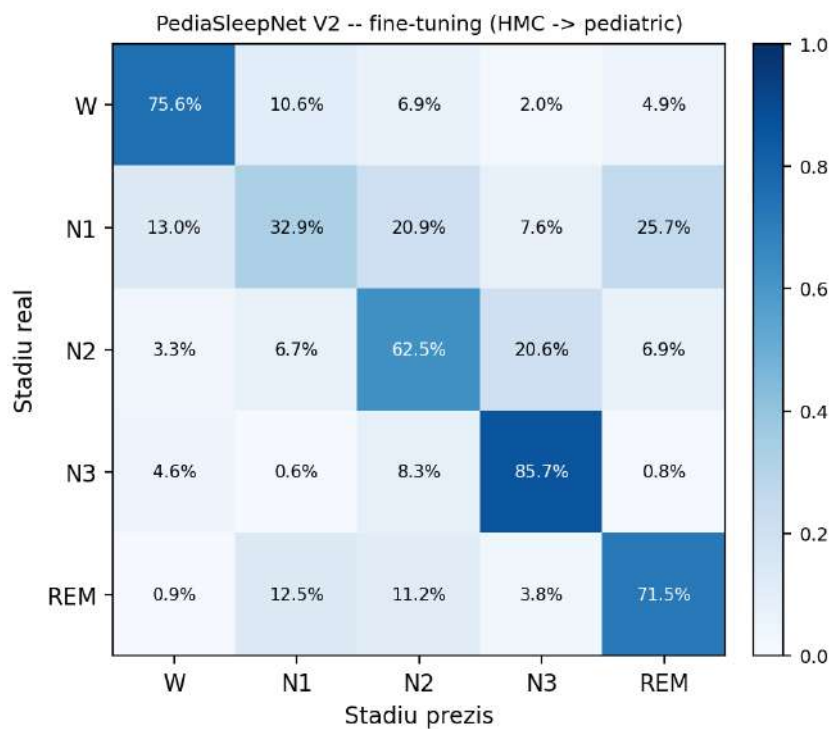


Figura 6.16 Matricea de confuzie normalizată – PediaSleepNet V2 (HMC - pediatric). Sursa: elaborare proprie.

Capitolul 7

Concluzii

Pipeline-ul experimental al acestei lucrări a urmat patru etape metodologice distincte. **(1) Preprocesarea semnalelor** a constat în citirea fișierelor EEG brute, conversia anotațiilor la schema de cinci clase AASM, segmentarea în epoci de 30 de secunde, reeșantionarea la rata țintă specifică fiecărui model și salvarea datelor preprocesate. **(2) Pre-antrenarea** a fost realizată pe seturile adulte Sleep-EDF și HMC, urmând protocolul original al fiecărei arhitecturi ($4 \times 2 = 8$ experimente). **(3) Adaptarea** modelelor pe setul pediatric s-a efectuat prin *transferul între domenii*, cu validare încrucișată pe 5 folduri (8 experimente). **(4) Evaluarea comparativă** a combinat acuratețea globală, *macro F1-score*, coeficientul Cohen's κ , *F1-score*-ul per clasă și matricele de confuzie normalizate. Acest capitol adună, pe scurt, rezultatele prezentate în Capitolul 6 și formulează concluziile principale ale lucrării.

7.1 Concluzii privind pre-antrenarea pe date adulte

În etapa de pre-antrenare, modelele de referință s-au comportat conform așteptărilor din literatură. **AttnSleep** a obținut cele mai bune metrice globale pe ambele seturi adulte (Sleep-EDF: ACC 95,5%, MF1 80,5%; HMC: ACC 85,3%, MF1 78,1%), iar **DeepSleepNet** a fost competitiv pe Sleep-EDF (ACC 89,3%, MF1 74,0%), dar instabil și slab pe HMC (ACC 55,5%, MF1 54,6%), din cauza populației clinice heterogene și a sensibilității modelului la volumul și diversitatea datelor. Sinteza tuturor celor 8 experimente de pre-antrenare este prezentată în Tabelul 7.1.

Model	Set	ACC	MF1	κ	N1 (F1)
DeepSleepNet	Sleep-EDF	89,3	74,0	0,80	36,4
AttnSleep	Sleep-EDF	95,5	80,5	0,89	42,7
PediaSleepNet Base	Sleep-EDF	87,2	75,0	0,77	42,2
PediaSleepNet V2	Sleep-EDF	86,9	74,5	0,77	43,4
DeepSleepNet	HMC	55,5	54,6	0,44	31,2
AttnSleep	HMC	85,3	78,1	0,80	44,6
PediaSleepNet Base	HMC	71,7	70,3	0,64	46,1
PediaSleepNet V2	HMC	72,3	69,5	0,64	40,9

Tabel 7.1 Sinteza rezultatelor de pre-antrenare pe seturile adulte (valori în %, κ adimensional). Cele mai bune valori pe fiecare set sunt îngroșate.

Rezultatele obținute pe Sleep-EDF sunt consistente cu valorile raportate în lucrările originale (Tabelul 7.2). DeepSleepNet original raportează pe canalul Fpz-Cz, în validare încrucișată pe 20 de folduri, ACC 82,0%, MF1 76,9 și $\kappa = 0,76$ [25]; AttnSleep raportează pe Sleep-EDF-20 ACC 84,4%, MF1 78,1, $\kappa = 0,79$, respectiv pe Sleep-EDF-78 ACC 81,3%, MF1 75,1, $\kappa = 0,74$ [6]. Valorile de acuratețe le ating sau le depășesc, diferențele provenind în principal din protocolul de partiționare diferit (un număr mai redus de folduri și un set de test mai mic, ceea ce poate produce estimări mai optimiste). Pentru HMC, set mai recent, nu există valori de referință în lucrările originale, motiv pentru care comparația cu literatura se limitează la Sleep-EDF.

Model	Sursă	ACC	MF1	κ	N1 (F1)
DeepSleepNet	Supratak et al. [25]	82,0	76,9	0,76	46,6
DeepSleepNet	lucrarea de față	89,3	74,0	0,80	36,4
AttnSleep	Eldele et al. (EDF-20) [6]	84,4	78,1	0,79	42,8
AttnSleep	lucrarea de față	95,5	80,5	0,89	42,7

Tabel 7.2 Comparație cu rezultatele raportate în literatură pe Sleep-EDF (valori în %).

Concluzia acestei etape este că, pe domeniul adult pe care au fost concepute, modelele de referință (în special AttnSleep) rămân cele mai performante, iar arhitecturile PediaSleepNet, deși competitive, nu aduc un avantaj global pe date adulte, ceea ce este de așteptat, întrucât mecanismele lor de condiționare sunt orientate spre semnalul pediatric.

7.2 Concluzii privind transferul pe date pediatrice

Etapă de *adaptare* pe setul pediatric, scopul central al lucrării, inversează ierarhia observată la pre-antrenare. **Modelele PediaSleepNet propuse depășesc clar modelele de referință pe date de copii**, indiferent de backbone-ul de pre-antrenare (Tabelul 7.3). Pe backbone-ul HMC, PediaSleepNet Base obține cel mai bun *macro F1-score* din întreaga lucrare pe setul pediatric (MF1 64,6%, ACC 70,1%), iar pe backbone-ul Sleep-EDF, PediaSleepNet V2 obține cea mai bună acuratețe (ACC 71,2%, MF1 64,1%). Prin contrast, DeepSleepNet se prăbușește pe date pediatrice (MF1 37,1–45,3%), iar AttnSleep, deși cel mai bun model de referință după transfer (MF1 până la 56,6%), rămâne cu 8–11 puncte de MF1 sub modelele PediaSleepNet.

Model	Backbone	ACC	MF1	κ	N1 (F1)
DeepSleepNet	Sleep-EDF	53,9	37,1	0,32	7,5
AttnSleep	Sleep-EDF	64,1	53,4	0,48	13,4
PediaSleepNet Base	Sleep-EDF	68,8	61,6	0,57	26,6
PediaSleepNet V2	Sleep-EDF	71,2	64,1	0,59	30,5
DeepSleepNet	HMC	58,4	45,3	0,39	12,3
AttnSleep	HMC	67,7	56,6	0,53	13,1
PediaSleepNet Base	HMC	70,1	64,6	0,59	29,2
PediaSleepNet V2	HMC	67,7	61,9	0,56	28,2

Tabel 7.3 Sinteza rezultatelor de *adaptare* pe setul pediatric (medie pe 5 folduri, valori în %). Cele mai bune valori pe fiecare backbone sunt îngroșate.

Acest rezultat confirmă ipoteza centrală a lucrării: **adaptarea explicită la specificul neurofiziologic al EEG-ului pediatric, prin filtrarea adaptivă depthwise pe grupe de vârstă și condiționarea SE–FiLM, produce un avantaj măsurabil pe populația țintă**, în ciuda numărului mult mai mic de parametri față de DeepSleepNet. Diferența este cea mai pronunțată exact acolo unde contează clinic: detecția stadiilor minoritare și a tranzițiilor specifice copiilor.

7.3 Limitele cercetării

Setul de date pediatric utilizat este redus ca volum ($N = 19$ pacienți, vârste 9–17 ani), ceea ce limitează puterea statistică a estimărilor și face imposibilă antrenarea *de novo* a arhitecturilor de *deep learning* cu milioane de parametri; *transferul între domenii* de pe date adulte a reprezentat singura abordare fezabilă în aceste condiții.

Resursele computaționale disponibile, un singur GPU NVIDIA H100 pe Google Colab, cu limitări de durată a sesiunilor, au restricționat numărul de iterații de căutare a hiperparametrilor, numărul de epoci de antrenare și posibilitatea de a rula studii de ablație la scară mai mare.

Toate modelele au fost antrenate și evaluate pe un singur canal EEG (F4-M1 pentru HMC, Fpz-Cz pentru Sleep-EDF și F4-C4 pentru setul pediatric). Această alegere limitează generalizarea: centrele clinice folosesc montaje diferite (de exemplu C4-M1, O2-M1, C3-M2), iar modelele nu au fost validate pentru funcționalitatea inter-montaj.

7.4 Contribuții proprii

Această secțiune sintetizează contribuțiile personale aduse în cadrul prezentei lucrări. Contribuțiile personale constau în:

- identificarea, selecția și analiza literaturii de specialitate relevante;
- selecția arhitecturilor de rețele neuronale adecvate pentru problema studiată;
- proiectarea și implementarea celor două modele PediaSleepNet (Base și V2);
- preprocesarea seturilor de date utilizate.
- implementarea și antrenarea modelelor pe seturi de date publice;
- adaptarea modelelor pentru populația pediatrică vizată prin tehnici de adaptare;
- compararea performanțelor obținute în vederea identificării celui mai precis model de clasificare;

7.5 Direcții viitoare de cercetare

Îmbunătățirea clasificării stadiului N1 rămâne cea mai urgentă problemă deschisă. Cercetările viitoare ar trebui să exploreze capete auxiliare mai sofisticate pentru detecția N1, funcții de pierdere mai complexe precum *focal loss* [18], modelarea bidirecțională a tranzițiilor inter-stadii, precum și integrarea unor modalități suplimentare (EOG, EMG), care furnizează informație discriminativă la granița veghe-somn, acolo unde EEG-ul singur este ambiguu.

Adaptabilitatea inter-montaj reprezintă o altă direcție esențială: arhitecturile viitoare ar trebui proiectate sau adaptate pentru a funcționa robust pe montaje EEG

diferite (F4-M1, C4-M1, O2-M1, C3-M2 etc.), permițând implementarea în contexte clinice diverse fără reantrenare completă de la zero.

Dacă această cercetare ar fi reluată de la început, aş investi semnificativ mai mult timp în explorarea literaturii de specialitate privind modelele de detecție automata pe pacienți pediatrici existente, pentru a fundamenta mai bine alegerile arhitecturale și a evita reinventarea unor soluții deja studiate. De asemenea, aş acorda o atenție sporită înțelegerii și îmbunătățirii detecției stadiului N1, atât prin analiza mai profundă a graniței neurofiziologice dintre N1, W și REM, cât și prin explorarea unor mecanisme de detecție specializate înainte de proiectarea arhitecturii complete.

În concluzie, lucrarea demonstrează că modelele de referință performante pe adulți nu se transferă direct la copii, iar arhitecturile PediaSleepNet propuse, concepute în jurul caracteristicilor EEG-ului de dezvoltare, oferă cel mai bun compromis între acuratețe, robustețe și detecția stadiilor dificile pe populația pediatrică, obiectivul central al acestei lucrări.

Bibliografie

- [1] Diego Alvarez-Estevez and Roselyne M. Rijsman. Haaglanden medisch centrum sleep staging database (version 1.1). *PhysioNet*, 2022. doi: 10.13026/t79q-fr32. Version 1.1.
- [2] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E. Hinton. Layer normalization. *arXiv preprint arXiv:1607.06450*, 2016.
- [3] Mark F. Bear, Barry W. Connors, and Michael A. Paradiso. *Neuroscience: Exploring the Brain*. Wolters Kluwer, Philadelphia, PA, 4th edition, 2015.
- [4] François Chollet. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1251–1258, 2017. doi: 10.1109/CVPR.2017.195.
- [5] Jacob Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46, 1960. doi: 10.1177/001316446002000104.
- [6] Emadeldeen Eldele, Zhenghua Chen, Chengyu Liu, Min Wu, Chee-Keong Kwoh, Xiaoli Li, and Cuntai Guan. An attention-based deep learning approach for sleep stage classification with single-channel EEG. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:809–820, 2021. doi: 10.1109/TNSRE.2021.3076234.
- [7] Ary L. Goldberger, Luis A. N. Amaral, Leon Glass, Jeffrey M. Hausdorff, Plamen Ch. Ivanov, Roger G. Mark, Joseph E. Mietus, George B. Moody, Chung-Kang Peng, and H. Eugene Stanley. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000. doi: 10.1161/01.CIR.101.23.e215.
- [8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016. doi: 10.1109/CVPR.2016.90.
- [9] Dan Hendrycks and Kevin Gimpel. Gaussian error linear units (GELUs). *arXiv preprint arXiv:1606.08415*, 2016.
- [10] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997. doi: 10.1162/neco.1997.9.8.1735.
- [11] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7132–7141, 2018. doi: 10.1109/CVPR.2018.00745.

- [12] Conrad Iber, Sonia Ancoli-Israel, Andrew Chesson, and Stuart F. Quan. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. American Academy of Sleep Medicine, Westchester, IL, 1st edition, 2007.
- [13] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, pages 448–456, 2015.
- [14] Bob Kemp and Joan Olivan. European data format ‘plus’ (EDF+), an EDF alike standard format for the exchange of physiological data. *Clinical Neurophysiology*, 114(9):1755–1761, 2003. doi: 10.1016/S1388-2457(03)00123-8.
- [15] Bob Kemp, Aeilko H. Zwinderman, Bert Tuk, Hilbert A. C. Kamphuisen, and Josefien J. L. Oberyé. Analysis of a sleep-dependent neuronal feedback loop: The slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering*, 47(9): 1185–1194, 2000. doi: 10.1109/10.867928.
- [16] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2015.
- [17] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 1097–1105, 2012.
- [18] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 2980–2988, 2017. doi: 10.1109/ICCV.2017.324.
- [19] Vinod Nair and Geoffrey E. Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th International Conference on Machine Learning (ICML)*, pages 807–814, 2010.
- [20] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. PyTorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 8024–8035, 2019.
- [21] Ethan Perez, Florian Strub, Harm de Vries, Vincent Dumoulin, and Aaron Courville. FiLM: Visual reasoning with a general conditioning layer. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), 2018.

- [22] Allan Rechtschaffen and Anthony Kales. *A Manual of Standardized Terminology, Techniques and Scoring Systems for Sleep Stages of Human Subjects*. UCLA Brain Information Service / Brain Research Institute, Los Angeles, CA, 1968.
- [23] Mike Schuster and Kuldip K. Paliwal. Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 45(11):2673–2681, 1997. doi: 10.1109/78.650093.
- [24] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15:1929–1958, 2014.
- [25] Akara Supratak, Hao Dong, Chao Wu, and Yike Guo. DeepSleepNet: A model for automatic sleep stage scoring based on raw single-channel EEG. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(11):1998–2008, 2017. doi: 10.1109/TNSRE.2017.2721116.
- [26] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 5998–6008, 2017.

Capitolul A

Anexe

Această anexă conține două părți: (1) configurația și hiperparametrii modelelor PediaSleepNet, și (2) vizualizări suplimentare ale setului de date pediatric descris în secțiunea 4.3.

A.1 Configurația și hiperparametrii modelelor

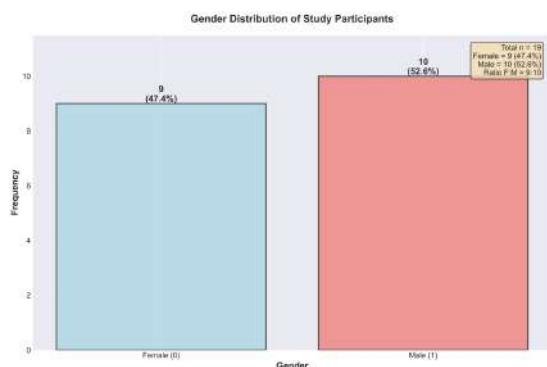
Tabelul A.1 sistematizează principalii hiperparametri ai celor două versiuni.

Parametru	PediaSleepNet Base	PediaSleepNet V2
Rată de eșantionare F_s	256 Hz	256 Hz
Canale de intrare	1 (EEG raw)	1 (EEG raw)
Grupe demografice	2 ($G0$, $G1$)	2 ($G0$, $G1$)
Dimensiune nucleu filtru adaptiv	33	33
Nuclee ramuri MRCNN+	50, 150, 400	50, 150, 400
Metodă fuziune MRCNN+	Concatenare temporală	Medie element-wise
d_{model}	128	128
Straturi Transformer în TCE	2	2
Capete de atenție în TCE	4	4
Context maxim	20 epoci	20 epoci
<i>Dropout</i> (MRCNN / TCE)	0,25 / 0,50	0,25 / 0,50
Cap auxiliar N1	Nu	Da (pierdere BCE)
Volum parametri	$\sim 1,1$ M	~ 945 K

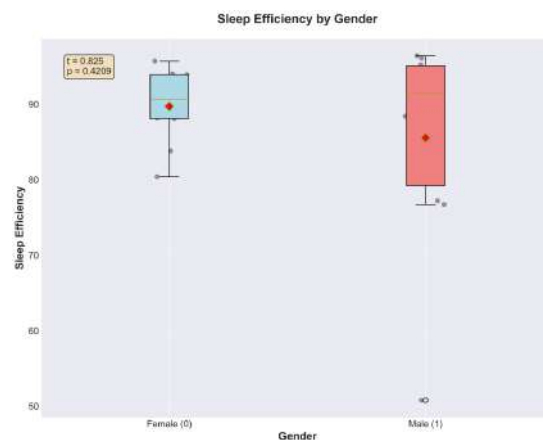
Tabel A.1 Hiperparametrii principali ai arhitecturilor PediaSleepNet Base și V2.

A.2 Setul de date pediatric

Graficele de mai jos completează analiza descriptivă din textul principal, oferind perspective adiționale asupra distribuției stadiilor de somn, eficienței somnului și variabilității inter-individuale în populația pediatrică ($N = 19$ pacienți, vârste 9–17 ani).

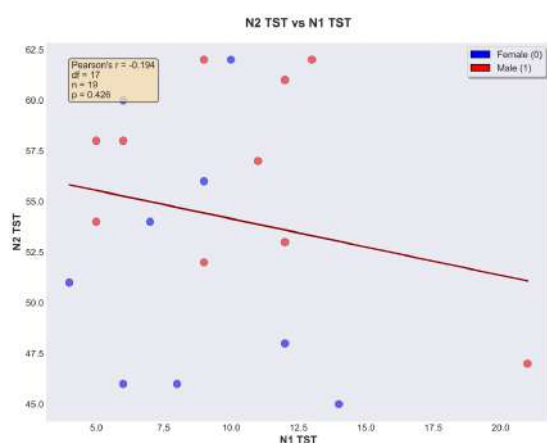


(a) Distribuția pe gen a pacienților pediatrici (10 băieți, 9 fete).

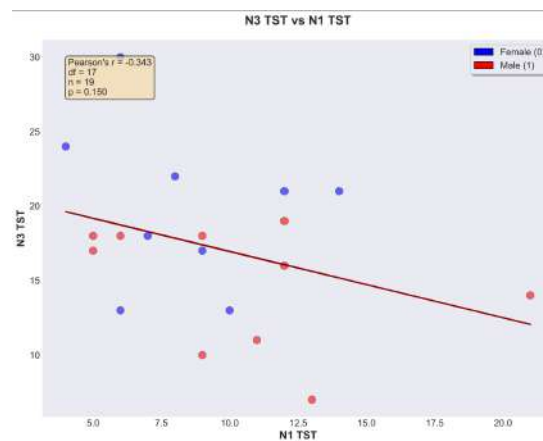


(b) Eficiența somnului în funcție de gen. Distribuțiile sunt similare între băieți și fete.

Figura A.1 Distribuția pe gen și eficiența somnului per gen în populația pediatrică. Sursa: elaborare proprie.

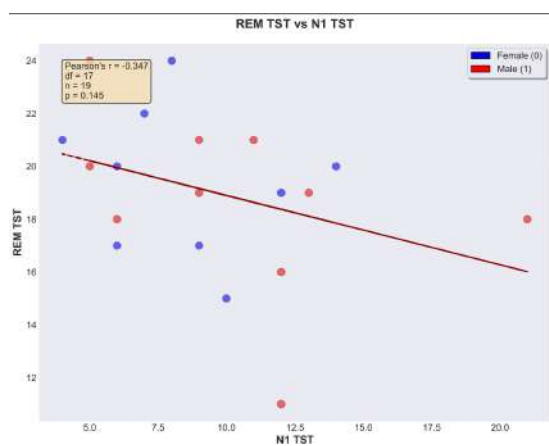


(a) Relația dintre proporțiile N1 și N2 din TST. N1 este rar (~9%) iar N2 este dominant (~54%).

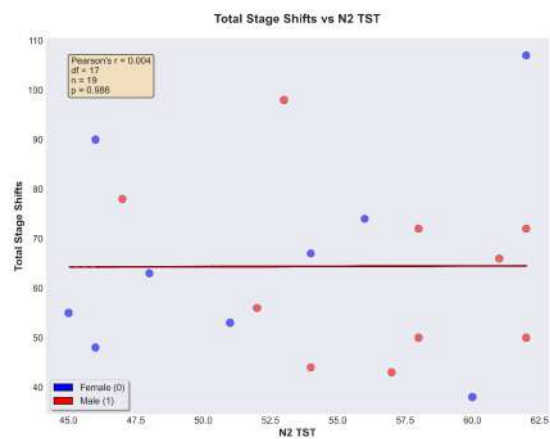


(b) Relația dintre proporțiile N1 și N3 din TST. Variabilitatea inter-individuală a N3 este ridicată (7–30%).

Figura A.2 Relațiile dintre proporțiile stadiilor NREM (N1 vs. N2 și N1 vs. N3). Sursa: elaborare proprie.



(a) Relația dintre N1 și REM din TST. Somnul REM variază între 11% și 24%.



(b) Relația dintre numărul de tranziții inter-stadii și proporția N2. pacienții cu mai multe tranziții tind să aibă un somn mai fragmentat.

Figura A.3 Relații între stadiile de somn și fragmentarea somnului în populația pediatrică. Sursa: elaborare proprie.